AM

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 January 2002 (31.01.2002)

PCT

(10) International Publication Number
WO 02/08447 A2

(51) International Patent Classification[7]: C12Q

(21) International Application Number: PCT/IL01/00669

(22) International Filing Date: 20 July 2001 (20.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/220,248 24 July 2000 (24.07.2000) US

(71) Applicant *(for all designated States except US)*: **INSTITUTE OF MOLECULAR & CELL BIOLOGY** [SG/SG]; 30 Medical Drive, 117609 Singapore (SG).

(71) Applicant *(for BB only)*: **EHRLICH, Gal** [IL/IL]; 28 Bezalel Street, 52 521 Ramat Gan (IL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TAN, Yin, Hwee** [SG/SG]; 3 Pandan Valley # 11-308, Chempaka Court, 597627 Singapore (SG). **LIM, Siew, Pheng** [SG/SG]; 427 #04-314 Serangoon Central, 550427 Singapore (SG). **LIM, Seng, Gee** [SG/SG]; 113 Clementi Road # 04-08, 129789 Singapore (SG). **HONG, Wan, Jin** [SG/SG]; 835 Bukit Timah Road, #01-07 Royal Ville, 279888 Singapore (SG).

(74) Agent: **G. E. EHRLICH (1995) LTD.**; 28 Bezalel Street, 52 521 Ramat Gan (IL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *entirely in electronic form (except for this front page) and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/08447 A2

(54) Title: NUCLEIC ACIDS AND METHODS FOR DETECTING VIRAL INFECTION, UNCOVERING ANTI-VIRAL DRUG CANDIDATES AND DETERMINING DRUG RESISTANCE OF VIRAL ISOLATES

(57) Abstract: A nucleic acid construct is provided. The nucleic acid construct includes (a) an expression cassette including: (i) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a core sequence of the RNA virus; (ii) a second polynucleotide region including a 3' UTR sequence of the RNA virus and at least a C-terminal portion of a polymerase sequence of the virus; and (iii) a third polynucleotide region encoding a reporter molecule, the third polynucleotide region being flanked by the first and the second polynucleotide regions; and (b) a promoter sequence being operatively linked to the expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette.

NUCLEIC ACIDS AND METHODS FOR DETECTING VIRAL INFECTION, UNCOVERING ANTI-VIRAL DRUG CANDIDATES AND DETERMINING DRUG RESISTANCE OF VIRAL ISOLATES

FIELD AND BACKGROUND OF THE INVENTION

The present invention relates to nucleic acid constructs and methods of utilizing same for detecting infection of an RNA virus, for uncovering anti-viral drug candidates and for determining drug resistance of isolates of an RNA virus. More particularly, the present invention relates to a nucleic acid construct which transcribes a minus strand RNA sequence encoding a reporter polypeptide and including 5' and 3' sequences of an RNA virus. When transcribed in a cell infected with an RNA virus capable of replicating the minus strand RNA sequence, a plus strand of this RNA sequence is formed and translated by the host cell into an active reporter polypeptide.

Viral diseases are some of the major scourges of mankind and include such virulent disorders as smallpox, yellow fever, rabies, poliomyelitis and AIDS. In addition, viruses carrying oncogenes are responsible for a number of human tumors and cancers.

It is a remarkable and proven fact that some virus infections occur without overt symptoms, while others can cause more than one clinical manifestation involving more than one organ system of the body. This lack of a defining clinical manifestation in some infections, presents a major hurdle to an accurate and timely diagnosis of infections, which in some cases is crucial for the prevention of disease and death.

Several diagnostic procedures have been developed in efforts to improve the detection and diagnosis of viral infections. These procedures involve the detection of viral components in cells of infected individuals or the detection of blood components generated as a response to the presence of a viral infection. Although such methods provide acceptable accuracy in

detecting some viral infections, they are oftentimes expensive and time consuming to carry out.

Although accurate and timely diagnosis of some viral infections provides clinicians with better chances of combating viral infection, the lack of suitable anti-viral drugs limits the possibilities of treatment for such viral infections

As such, for the past decades, universities and pharmaceutical companies have invested considerable resources in efforts to uncover potential anti-viral drug candidates and/or to determine the anti-viral drug resistance of some viruses.

Present day anti-viral drug screening methods rely on detecting interactions between viral components and molecules having potential anti-viral activity. For example, the identification of inhibitors of virally encoded proteases ("protease inhibitors") relies on the in-vitro screening of purified viral protease with chemical compounds in the presence of synthetic peptide substrates. Initial in-vitro screening is usually followed by a bioassay designed for determining whether a potential protease inhibitor or its derivatives function in virally infected cells prior to additional testing conducted in more complex biological systems.

Screening for drug resistance of certain virus isolates is typically effected by phenotypic testing (plaque reduction assay). This is a labor intensive, time consuming and expensive technique that oftentimes does not correlate well to the clinical response to drug therapy in individual patients. Nonetheless, because of its derivation from testing for sensitivity to antibacterial agents, this technique is often considered to be the "gold standard".

Prior art drug and drug resistance screening methods, such as the methods described above, are further limited in that such methods are not readily utilizable in screening for molecules possessing anti-viral activities

against, nor can they be utilized to determine the drug resistance of, RNA viruses.

A large portion of the viruses responsible for human diseases are RNA viruses. Since the RNA genome of such viruses is replicated via an RNA intermediate, recombinant manipulation thereof for the purposes of constructing cell, or cell free assays is oftentimes a difficult task. In addition, the high heterogeneity of RNA viral genomes further complicates recombinant manipulation and also limits the accuracy of prior art cell free drug and drug resistance screenings.

One example of a disease causing RNA virus is the Hepatitis C virus (HCV) which is a member of the Flaviviridae family, and the major cause of chronic liver disease worldwide (1, 2). HCV is an enveloped virus with a single-stranded, positive sense, RNA genome that encodes a single open reading frame (ORF) of about 3010 amino acids (aa) which is co-translationally and post-translationally cleaved to give rise to at least 10 polypeptides (3). Located at its N-terminal end are three structural proteins, followed by at least seven non-structural (NS) proteins (1). Combined action of host-derived signal peptidase(s) and the virus-encoded proteases are involved in the processing of this polyprotein (4-8).

Similar to other RNA viruses, the genome of HCV is highly heterogeneous, and several genotypes and subtypes have been described (12, 13). Numerous studies have successfully demonstrated partial replication of the virus in *in-vitro* culture systems using human T-cells, B-cells (9, 10), human hepatocytes (11, 12) or chimpanzee hepatocytes (13, 14). However, these systems suffer from low viral replication efficiency and limited passage cycles. More recently, high level replication of subgenomic HCV RNA was established in a human hepatoma cell line that would enable long-term production of viral RNA and proteins (14). Unfortunately, the complete life cycle of virus does not take place in this

system nor are infectable virions produced as transfection with the full length genome failed to produce any viable cell clones (14).

Replication of HCV *in vivo* involves the replication of its single positive-stranded RNA through negative (anti-sense) strand intermediates via the NS5B polymerase (15-17). The negative strand RNA formed then serves as a template for the synthesis of more positive RNA strands which are either used as templates for translation of viral proteins or packaged for production of viral particles. Binding and initiation of reverse strand synthesis by NS5B is dependent on stem-loop structures present in the 3' of the viral genome (17, 18). Based on this knowledge the inventors of the present invention decided to create a reporter system using constructs encoding anti-sense luciferase gene flanked by HCV 5' and 3' NCR.

While reducing the present invention to practice, a cDNA clone encoding a complete HCV genome was generated by the present inventors. Sequences derived from this cDNA clone were incorporated in novel chimeric HCV-luciferase expression constructs which can be used, according to the teachings of the present invention, in accurate and rapid cell based assays for detecting HCV infection, screening molecules for potential anti-viral activities and determining drug resistance of HCV isolates.

SUMMARY OF THE INVENTION

According to one aspect of the present invention there is provided a nucleic acid construct comprising: (a) an expression cassette including: (i) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of the RNA virus; (ii) a second polynucleotide region including a 3' UTR sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus; and (iii) a third polynucleotide region encoding a reporter molecule, the third polynucleotide region being flanked by the first and the second

polynucleotide regions; and (b) a promoter sequence being operatively linked to the expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette.

According to another aspect of the present invention there is provided a genetically transformed cell comprising a nucleic acid construct including: (a) an expression cassette including: (i) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of the RNA virus; (ii) a second polynucleotide region including a 3' UTR sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus; and (iii) a third polynucleotide region encoding a reporter molecule, the third polynucleotide region being flanked by the first and the second polynucleotide regions; and (b) a promoter sequence being operatively linked to the expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette.

According to further features in preferred embodiments of the invention described below, the genetically transformed cell further comprising an additional nucleic acid construct for expressing at least an RNA dependent RNA polymerase of a virus, whereas the first and the second polynucleotide regions being selected such that the RNA dependent RNA polymerase is capable of replicating the minus strand RNA molecule into plus strand RNA.

According to still further features in the described preferred embodiments at least a portion of the first polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 1-374 of SEQ ID NO:33.

According to still further features in the described preferred embodiments at least a portion of the second polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 9158-9609 of SEQ ID NO:33.

According to still further features in the described preferred embodiments the first polynucleotide region further includes a 5' UTR sequence of the RNA virus.

According to still further features in the described preferred embodiments the first polynucleotide region includes an IRES sequence.

According to still further features in the described preferred embodiments the RNA virus is selected from the group consisting of a positive strand RNA virus and a negative strand RNA virus.

According to still further features in the described preferred embodiments the RNA virus is selected from the group consisting of a virus of the picornavirus family, a virus of the togavirus family, a virus of the orthomyxovirus family, a virus of the paramyxovirus family, a virus of the coronavirus family, a virus of the calicivirus family, a virus of the arenavirus family, a virus of the rhabdovirus family and a virus of the bunyavirus family.

According to still further features in the described preferred embodiments the RNA virus is Hepatitis C.

According to still further features in the described preferred embodiments the first and the second polynucleotide regions are selected such that the minus strand RNA molecule transcribable from the expression cassette is replicatable by an RNA dependent RNA polymerase of the virus into a plus strand RNA molecule.

According to still further features in the described preferred embodiments the promoter is functional in a eukaryotic cell.

According to still further features in the described preferred embodiments the eukaryotic cell is selected from the group consisting of an insect cell, a yeast cell and a mammalian cell.

According to still further features in the described preferred embodiments the reporter molecule is a polypeptide selected from the group consisting of an enzyme, a fluorophore, a substrate and a ligand.

According to yet another aspect of the present invention there is provided a method of detecting a presence of an RNA virus in a cell, the method comprising the steps of: (a) incubating a nucleic acid construct with an extract of the cell under conditions suitable for transcription and translation of the nucleic acid construct, the nucleic acid construct including: (i) an expression cassette having: (one) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of the RNA virus; (two) a second polynucleotide region including a 3' UTR sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus; and (three) a third polynucleotide region encoding a reporter molecule, the third polynucleotide region being flanked by the first and the second polynucleotide regions; and (ii) a promoter sequence being operatively linked to the expression cassette in a manner so as to direct the transcription of a minus strand RNA molecule from the expression cassette when the nucleic acid construct is incubated with the extract, the first and the second polynucleotide regions being selected such that the minus strand RNA molecule transcribed is replicatable by the polymerase of the RNA virus into a plus strand RNA molecule; and (b) quantifying a level of the reporter molecule to thereby determine the presence of the virus in the cell.

According to still further features in the described preferred embodiments the reporter molecule is a polypeptide translated from the plus strand RNA molecule.

According to still further features in the described preferred embodiments the method described above further comprising the step of comparing the level of the reporter molecule to that obtained from cells free of the virus.

According to a further aspect of the present invention there is provided a method of screening for anti-viral drugs, the method comprising the steps of: (a) co-incubating a nucleic acid construct, a polynucleotide

encoding at least a polymerase of an RNA virus and a potential anti-viral molecule under conditions suitable for transcription and translation of the nucleic acid construct and the polynucleotide encoding at least the polymerase, the nucleic acid construct including: (i) an expression cassette having: (one) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of the RNA virus; (two) a second polynucleotide region including a 3' UTR sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus; and

(three) a third polynucleotide region encoding a reporter molecule, the third polynucleotide region being flanked by the first and the second polynucleotide regions; and (ii) a promoter sequence being operatively linked to the expression cassette in a manner so as to direct the transcription of a minus strand RNA molecule from the expression cassette when the nucleic acid construct is incubated with the polynucleotide encoding the polymerase of the RNA virus under the conditions suitable for transcription and translation, the first and the second polynucleotide regions being selected such that the minus strand RNA molecule transcribed is replicatable by the polymerase of the RNA virus into a plus strand RNA molecule; and (b) quantifying a level of the reporter molecule to thereby determine the anti-viral activity of the potential anti-viral molecule.

According to still further features in the described preferred embodiments the reporter molecule is a polypeptide translated from the plus strand RNA molecule.

According to still further features in the described preferred embodiments the method described above further comprising the step of comparing the level of the reporter molecule to that obtained from cells free of the virus.

According to still further features in the described preferred embodiments the potential anti-viral molecule is selected from the group

consisting of a nucleoside or nucleotide analogue and an immune-modulatory molecule.

According to still further features in the described preferred embodiments step (a) is effected by introducing the nucleic acid construct, the polynucleotide encoding at least the polymerase of the RNA virus and the potential anti-viral molecule into a cell.

According to still further features in the described preferred embodiments step (a) is effected by introducing the nucleic acid construct and the potential anti-viral molecule into a cell infected with the RNA virus.

According to yet a further aspect of the present invention there is provided a method of determining drug resistance of an RNA virus, the method comprising the steps of: (a) co-incubating a nucleic acid construct, a polynucleotide encoding at least a polymerase of the RNA virus and an anti-viral drug molecule under conditions suitable for transcription and translation of the nucleic acid construct and the polynucleotide encoding at least the polymerase, the nucleic acid construct including: (i) an expression cassette having: (one) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of the RNA virus; (two) a second polynucleotide region including a 3' UTR sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus; and (three) a third polynucleotide region encoding a reporter molecule, the third polynucleotide region being flanked by the first and the second polynucleotide regions; and (ii) a promoter sequence being operatively linked to the expression cassette in a manner so as to direct the transcription of a minus strand RNA molecule from the expression cassette when the nucleic acid construct is incubated with the polynucleotide encoding at least the polymerase of the RNA virus under the conditions suitable for transcription and translation, the first and the second polynucleotide regions being selected such that the minus strand RNA molecule transcribed is replicatable by the polymerase of the RNA virus

into a plus strand RNA molecule; and (b) quantifying a level of the reporter molecule to thereby determine the resistance of the RNA virus to the anti-viral drug.

According to still further features in the described preferred embodiments the method described above further comprising the step of comparing the level of the reporter molecule to that obtained from cells free of the anti-viral drug.

According to still further features in the described preferred embodiments the reporter molecule is a polypeptide translated from the plus strand RNA molecule.

According to still further features in the described preferred embodiments the anti-viral drug is selected from the group consisting of a nucleoside or nucleotide analogue and an immune-modulatory molecule.

According to still further features in the described preferred embodiments step (a) is effected by introducing the nucleic acid construct, the polynucleotide encoding at least the polymerase of the RNA virus and the anti-viral drug into a cell.

According to still further features in the described preferred embodiments step (a) is effected by introducing the nucleic acid construct and the anti-viral drug into a cell infected with the RNA virus.

The present invention successfully addresses the shortcomings of the presently known configurations by providing nucleic acid constructs and methods of utilizing same for detecting the presence of an RNA virus in a cell or a cell extract, for uncovering novel anti-viral drugs and for determining the resistance of RNA virus isolates to anti-viral drugs.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention is herein described, by way of example only, with reference to the accompanying drawings. With specific reference now to

the drawings in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only, and are presented in the cause of providing what is believed to be the most useful and readily understood description of the principles and conceptual aspects of the invention. In this regard, no attempt is made to show structural details of the invention in more detail than is necessary for a fundamental understanding of the invention, the description taken with the drawings making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

In the drawings:

FIG. 1A is a schematic representation of the overlapping HCV cDNA clones of HCV-S1 utilized in constructing the HCV genome. The positions of the first and last nucleotides and amino acids of the individual HCV proteins as well as the first and last nucleotide of the HCV 5' UTR and 3' UTR are indicated. Clones A-M represent the overlapping cDNA clones of HCV-S1 obtained from RT-PCR. The first and last nucleotide of each clone is indicated.

FIG. 1B illustrates the step employed for constructing the sense and antisense chimeric vectors of the present invention.

FIGs. 2A-C illustrate the protein products of *in vitro* translation experiments of HCV constructs separated on SDS-PAGE. Figure 2A - translation of the entire non-structural HCV polyprotein from pcDNA3(NSP). Figure 2B - translation of the entire structural HCV polyprotein from pcDNA(SP). Figure 2C - translation of the full length HCV genome from pcDNA3(S1). CPMM represents incubation with canine pancreatic microsomal membranes. Arrows indicate positions of autolytically cleaved products upon prolonged incubation. Molecular weight marker sizes (in kDa) are indicated on the left.

FIGs. 3A-G illustrate western analysis of 293T cells transiently transfected with pXJ41(S1). Cells were harvested two days post-transfection and lysate proteins were separated on SDS-PAGE gels and transferred onto nitrocellulose membranes. The blotted proteins were probed with anti-E2 (Figure 3A), anti-NS3 (Figure 3B) and anti-NS5A (Figure 3C) monoclonal antibodies. The detection of core (Figures 3E-G) and NS5B (Figures 3D-E) proteins, was effected using different sera from HCV infected patient at a dilution of 1:100 (Figures 3D-G). The immunoblots of Figures 3D-E represent sera taken from the same patient from which the HCV-S1 was cloned. Molecular weight marker sizes (in kDa) are indicated on the left.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is of nucleic acid constructs and methods utilizing same which can be utilized for detecting infection of an RNA virus, for uncovering anti-viral drug candidates and for determining drug resistance of isolates of an RNA virus. Specifically, the present invention is of a nucleic acid construct which transcribes a minus strand RNA sequence encoding a reporter polypeptide and including 5' and 3' sequences of an RNA virus. When transcribed in a cell infected with an RNA virus capable of replicating the minus strand RNA sequence, a plus strand of this RNA sequence is formed and translated by the host cell into an active reporter polypeptide.

The principles and operation of the present invention may be better understood with reference to the drawings and accompanying descriptions.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and the arrangement of the components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments or of being practiced or carried out in various

ways. Also, it is to be understood that the phraseology and terminology employed herein is for the purpose of description and should not be regarded as limiting.

The molecular studies of the pathogenesis of HCV and the development of anti-viral drugs have been hampered in part by the lack of a robust, cell-based assay to monitor viral replication. The currently available cell-based systems are limited by the low viral replication efficiency and limited passage cycles. Although high levels of replication of subgenomic HCV RNA was established in a human hepatoma cell line that would enable long-term production of viral RNA and proteins, this does not truly measure viral replication. The complete life cycle of HCV does not occur in this system, nor are infectable virions produced. Moreover, the authors failed to generate any viable cell clones when they carried out transfections with the full length genome (14).

Replication of the HCV genome *in vivo* is dependent in part on the proteolytic activity of host signal peptidase(s) for cleavage of its structural genes and on its NS3 protein, which systematically cleaves the viral NS polyprotein to release the individual active subunits (7). Of these, the viral RNA dependent RNA polymerase, NS5B, plays a vital role in replication through synthesis of both positive and negative viral RNA strands (15). Due to the low replication efficiency of HCV, nested RT-PCR for amplifying minus-strand RNA is employed to determine viral replication *in vivo*. This method is both laborious and easily prone to false positive errors. Although its sensitivity and reliability has been improved with the use of tagged primers and Tth polymerase (13), it still remains expensive and time-consuming.

As is further described in the Examples section which follows, to generate a reliable and simple reporter assay system which can be utilized to detect hepatitis C virus (HCV) replication *in vivo*, and to uncover novel anti-viral drugs as well as to screen for drug resistance in viral isolates, the

present inventors undertook the laborious task of generating a replication-competent full length HCV genome.

Sequences derived from this clone were then utilized to generate reporter expression constructs which produce a reporter signal in the presence of infecting virus particles.

Thus, according to one aspect of the present invention there is provided a nucleic acid construct. The nucleic acid construct includes an expression cassette having a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of the RNA virus, such as for example the N-terminal portion of the core sequence, and a second polynucleotide region including a 3' UTR sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the virus, such as for example a C-terminal portion of the viral polymerase sequence. The expression cassette also includes a third polynucleotide region which encodes a reporter polypeptide such as for example, an enzyme, a substrate, a ligand or receptor or a fluorophore.

According to the present invention, the reporter molecule encoding region is flanked by the first and the second polynucleotide regions and is in transcriptional linkage therewith.

The nucleic acid construct according to this aspect of the present invention, also includes a promoter sequence which serves to direct transcription of the expression cassette sequence in eukaryotic cells such as for example, mammalian cells, yeast cells or insect cells.

The promoter sequence is oriented with respect to the expression cassette sequence, such that transcription therefrom generates a minus strand RNA molecule.

As used herein the phrase "minus (or negative) strand RNA" refers to the complementary RNA strand of the "plus (or positive) strand RNA" which is the strand typically translated by the ribosomes into a polypeptide sequence.

According to a preferred embodiment of the present invention, at least a portion of the first polynucleotide region is at least 50 %, at least 60 %, at least 70 % at least 80 %, at least 90 to 95 % identical to a sequence encompassed by nucleotides 1-374 of SEQ ID NO:33.

According to another preferred embodiment of the present invention, at least a portion of the second polynucleotide region is at least 50 %, at least 60 %, at least 70 % at least 80 %, at least 90 to 95 % identical to a sequence encompassed by nucleotides 9158-9609 of SEQ ID NO:33.

Since the nucleic acid construct of the present invention transcribes a minus strand RNA molecule in cells, such a construct cannot generate an active reporter molecule in cells transformed with this construct. However, in the presence of a viral polymerase, such as the RNA dependent RNA polymerase encoded by RNA viruses (hereinafter RNA polymerase), such as the case when the transformed cell is infected with a virus or expresses the viral polymerase, replication of the minus strand RNA takes place and a plus strand RNA molecule is formed. This molecule can then be translated by the host cell ribosome into an active reporter molecule. It will be appreciated that this is true only in cases where the viral RNA polymerase binds and initiates replication from the viral sequences included within the transcribed minus strand RNA. In most cases, the viral sequences utilized in the expression cassette of the nucleic acid construct will be derived from the virus of interest, although in some cases, RNA polymerases of one virus can replicate RNA which includes 5' and 3' sequences from another virus.

Since the sequences regulating RNA replication in RNA viruses reside in the 5' and 3' NCRs and/or UTRs, such sequences alone are often sufficient in promoting RNA replication of the minus strand RNA transcribed from the nucleic acid construct of the present invention. However, not withstanding from the above, in some RNA viruses, coding region sequences are often necessary in order to initiate or enhance replication, as is the case for HCV. As such, the expression cassette

according to the present invention preferably also includes such sequences, the identity thereof can be determined by quantifying replication from various expression cassettes which include different segments from the coding region of the virus.

Since the cap dependent translation of RNA in virally infected cells is oftentimes downregulated by the presence of a replicating virus, the expression cassette preferably also include internal ribosome entry site (IRES) sequences for initiation of cap independent translation of the chimeric polypeptide(s)if such sequences are not already included within the 5' and 3' sequences.

The viral sequences included in the expression cassette according to the present invention, are derived from a plus strand RNA virus or a minus strand RNA virus such as for example a virus of the picornavirus family, a virus of the togavirus family, a virus of the orthomyxovirus family, a virus of the paramyxovirus family, a virus of the coronavirus family, a virus of the calicivirus family, a virus of the arenavirus family, a virus of the rhabdovirus family or a virus of the bunyavirus family.

According to another preferred embodiments of the present invention, the RNA virus is a Hepatitis C virus (HCV).

The nucleic acid construct described hereinabove can be constructed using commercially available mammalian expression vectors or derivatives thereof. Examples of suitable vectors include, but are not limited to, pcDNA3, pcDNA3.1(+/-), pZeoSV2(+/-), pSecTag2, pDisplay, pEF/myc/cyto, pCMV/myc/cyto, pCR3.1, which are available from Invitrogen, pCI which is available from Promega, pBK-RSV and pBK-CMV which are available from Stratagene, pTRES which is available from Clontech, and their derivatives and modificants.

Any of the promoter and/or regulatory sequences included in the mammalian expression vectors described above can be utilized to direct the transcription of the expression cassettes described above. However, since

such vectors are readily amenable to sequence modifications via standard recombinant techniques, additional regulatory elements, promoter and/or selection markers can easily be incorporated therein if needed.

The nucleic acid construct according to this aspect of the present invention can be utilized in a cell-based or a cell free assay to detect virus infection of a cell, to uncover novel anti-viral drugs or to determine the resistance of an RNA virus isolate to anti-viral drugs.

When utilized in cell-based assays, the nucleic acid construct is introduced into a cell via any standard transformation method. Numerous methods are known in the art for introducing exogenous polynucleotide sequences into eukaryotic cells. Such methods include, but are not limited to, direct polynucleotide uptake techniques, and virus or liposome mediated transformation (for further detail see, for example, "Methods in Enzymology" Vol. 1-317, Academic Press). Bombardment of cells or cell cultures.

A genetically transformed cell including the nucleic acid construct of the present invention either stability integrated into it's genome, or transiently expressed can be utilized for a cell-based assay. In assays designed for uncovering novel anti-viral drugs or determining the resistance of an RNA virus isolate to anti-viral drugs, such a cell can further be genetically transformed to also express an RNA polymerase of a virus of interest along with other viral proteins and as such serve as a "test bed" for various molecules of interest.

Thus, the nucleic acid construct of the present invention can be utilized in a method for detecting a presence of an RNA virus in a cell by incubating the nucleic acid construct with an extract of cell or by introducing the construct into the cell and measuring the signal from the reporter molecule. Preferably, this signal is compared to a signal measured from a cell infected with a virus and possibly also a cell not infected with the virus to thereby determine the presence of the virus in the cell.

As mentioned hereinabove, the nucleic acid construct of the present invention can be utilized in an assay designed for screening anti-viral activities of various molecules or in an assay for determining the drug resistance of an RNA virus isolate. Such assays are separately effected by incubating the nucleic acid construct and a potential anti-viral drug when screening molecules for anti-viral activities, or a known anti-viral drug when determining drug resistance of an RNA virus along with a cellular extract from an infected cell. Alternatively the constructs and potential or known drug are introduced into an infected cell or a cell expressing the viral polymerase and possibly other viral components.

Following a predetermined time period, the reporter activities are measured and preferably compared to those measured from cells not including the potential or known drug to thereby determine the anti-viral activity of the drug candidate or to determine the resistance of the virus to the known anti-viral drug.

It will be appreciated that although cell-free assays (*in-vitro*) can be efficiently utilized for determining the anti-viral activity of a drug candidate or for determining the resistance of the virus to the known anti-viral drug cell-based assays (*in-situ*) screening in virally infected cells is preferred since this method determines anti-viral activity *in-situ* and in the presence of all the virally expressed components and as such it is more accurate in predicting future activity of screened molecules *in-vivo*.

Thus, the present invention provides nucleic acid constructs and methods of utilizing same to detect viruses in infected cells, to screen and uncover potential anti-viral drugs and to determine drug resistance of virus isolates.

The present invention presents several advantages over prior art methods. It is easily to implementable and executable, and in addition when utilized for uncovering potential viral drugs and for drug resistance

screening it can provide results of an accuracy which far exceeds that achieved by presently available in-vitro methods.

Additional objects, advantages, and novel features of the present invention will become apparent to one ordinarily skilled in the art upon examination of the following examples, which are not intended to be limiting. Additionally, each of the various embodiments and aspects of the present invention as delineated hereinabove and as claimed in the claims section below finds experimental support in the following examples.

EXAMPLES

Reference is now made to the following examples, which together with the above descriptions, illustrate the invention in a non limiting fashion.

Generally, the nomenclature used herein and the laboratory procedures utilized in the present invention include molecular, biochemical, microbiological and recombinant DNA techniques. Such techniques are thoroughly explained in the literature. See, for example, "Molecular Cloning: A laboratory Manual" Sambrook et al., (1989); "Current Protocols in Molecular Biology" Volumes I-III Ausubel, R. M., ed. (1994); Ausubel et al., "Current Protocols in Molecular Biology", John Wiley and Sons, Baltimore, Maryland (1989); Perbal, "A Practical Guide to Molecular Cloning", John Wiley & Sons, New York (1988); Watson et al., "Recombinant DNA", Scientific American Books, New York; Birren et al. (eds) "Genome Analysis: A Laboratory Manual Series", Vols. 1-4, Cold Spring Harbor Laboratory Press, New York (1998); methodologies as set forth in U.S. Pat. Nos. 4,666,828; 4,683,202; 4,801,531; 5,192,659 and 5,272,057; "Cell Biology: A Laboratory Handbook", Volumes I-III Cellis, J. E., ed. (1994); "Current Protocols in Immunology" Volumes I-III Coligan J. E., ed. (1994); Stites et al. (eds), "Basic and Clinical Immunology" (8th Edition), Appleton & Lange, Norwalk, CT (1994); Mishell and Shiigi (eds), "Selected Methods in Cellular Immunology", W. H. Freeman and Co., New

York (1980); available immunoassays are extensively described in the patent and scientific literature, see, for example, U.S. Pat. Nos. 3,791,932; 3,839,153; 3,850,752; 3,850,578; 3,853,987; 3,867,517; 3,879,262; 3,901,654; 3,935,074; 3,984,533; 3,996,345; 4,034,074; 4,098,876; 4,879,219; 5,011,771 and 5,281,521; "Oligonucleotide Synthesis" Gait, M. J., ed. (1984); "Nucleic Acid Hybridization" Hames, B. D., and Higgins S. J., eds. (1985); "Transcription and Translation" Hames, B. D., and Higgins S. J., eds. (1984); "Animal Cell Culture" Freshney, R. I., ed. (1986); "Immobilized Cells and Enzymes" IRL Press, (1986); "A Practical Guide to Molecular Cloning" Perbal, B., (1984) and "Methods in Enzymology" Vol. 1-317, Academic Press; "PCR Protocols: A Guide To Methods And Applications", Academic Press, San Diego, CA (1990); Marshak et al., "Strategies for Protein Purification and Characterization - A Laboratory Course Manual" CSHL Press (1996); all of which are incorporated by reference as if fully set forth herein. Other general references are provided throughout this document. The procedures therein are believed to be well known in the art and are provided for the convenience of the reader. All the information contained therein is incorporated herein by reference.

EXAMPLE 1

MATERIALS AND METHODS

Clinical characteristics of the recipient patient:

Sera from an individual known to be suffering both thalessemia and chronic hepatitis C was used for RT-PCR to obtain overlapping clones comprising the full length HCV genome (SEQ ID NO:33). Sera was collected from the patient after bone marrow transplantation upon diagnosis of elevated levels of serum transaminase indicative of HCV reactivation. The patient was determined to be HCV positive by RT-PCR of the plasma using the branched DNA assay with a level of 35.2 Meq/ml (Quantiplex HCV RNA assay, version 2.0 (bDNA); Chiron Diagnostics). Serum samples were collected in 400 µl aliquots and stored at –80 °C.

Isolation of HCV RNA:

RNA was extracted from 400 µl of sera using 1.2 ml of the Trizol LS reagent (Gibco BRL, Gaithesburg MD, USA). The mixture was inverted for 20 seconds at room temperature (RT), 0.35 ml of chloroform was added and the mixture inverted again for 20 seconds. The mixture was allowed to stand at RT for 5 minutes following which it was centrifuged at 12 000 rpm for 20 minutes. The upper phase of the mixture was transferred to a new microfuge tube, 0.8 ml of isopropanol and mixing was effected via inversion. The tube was left at RT for 5 minutes following which it was spun again at 12 000 rpm for 20 minutes at 4 °C. The RNA pellet was air-dried and re-suspended in 50 ml of DEPC-treated water.

RT-PCR:

Several RT-PCR reactions were conducted in order to obtain the various overlapping cDNA fragments. The various RT-PCR utilized are listed in Table 1. The RNA, extracted as described above, was reverse transcribed at 42 °C for 1 hour using 100 ng of oligo(dT) and/or specific antisense primers and 200 U of Superscript II polymerase (Gibco BRL, Gaithersburg). The resultant cDNA samples were heated at 70 °C for 15 minutes and PCR amplified using the Expand High Fidelity PCR System (Boehringer Mannheim). The PCR reactions were performed with 2-5 µl of template in a total volume of 50 µl. Different cycling profiles were used depending on the target length and the melting temperature (Tm) of the primers. Generally the PCR conditions were as follows: a hot-start at 95 °C for 3 min, denaturation at 95 °C for 1 min, annealing at 45-65 °C for 1 min, and extension at 68 °C for 1 min per 1 kb of amplified cDNA. At the end of 30-35 cycles, a final extension was carried out at 68 °C for 8 minutes. In several cases nested PCR was carried out to obtain the HCV cDNA fragment (Table 1).

5' RACE:

To clone the 5' UTR of HCV, a 5' rapid amplification of cDNA ends method using the 5'/3' RACE kit from Boehringer Mannheim was employed. The first strand cDNA was synthesized with the antisense primer H3 (Table 1) and AM reverse transcriptase at 55 °C for 1 hour and the resultant cDNA was purified using the High Pure PCR Product Purification kit (Boehringer Mannheim). A terminal transferase was utilized for 3' dA-tailing of the purified cDNA sample following which the transferase was heat-inactivated at 70 °C for 10 minutes. The tailed cDNA was amplified using the oligo dT-anchor primer and the H29 and the gene specific H4 primers (Table 1) utilizing the Expand High Fidelity PCR system. PCR conditions were as follows: 95 °C for 3 minutes, followed by 35 cycles of 95 °C for 1 minute, 45 °C for 1 minute, 68 °C for 1 minute, and a final extension at 68 °C for 8 minutes. A second round of PCR was performed with 1 ml of the first reaction mixture and the PCR anchor primer and the H30 and H5 primers (Table 1). The PCR products were cloned into the pCRII TOPO plasmid using the TOPO TA cloning kit from Clontech (Carlsbad, CA, USA).

Construction of HCV-S1 cDNA clones encoding the structural proteins:

The region spanning the 5' non-coding region (NCR) including the p7 region (nucleotides -276 to 2461 in Figure 1A) was PCR amplified using clones C and D as templates and primers H2 and H12 (Table 1). The resulting 2.7 kb PCR product (nucleotides 65-2802 of SEQ ID NO:33) and a 600 bp PCR product comprising the NS2 cDNA (nucleotides 2769-3369 of SEQ ID NO:33) were used as templates for the H2 and H32 primers in a second round of PCR amplification (Table 1) to produce a 3.3 kb DNA fragment (nucleotides 65-3114 of SEQ ID NO:33). This PCR product and clone A were used as templates in a third round of PCR with primers H30 and H32. The resultant PCR product (nucleotides 1-3114 of SEQ ID NO:

33) was cloned into pXL TOPO TA vector from Clontech (Carlsbad, CA, USA) to generate clone J (Figure 1A). The truncated NS2 PCR product was amplified from clone E (Figure 1A) using the primers H31 and H32. The PCR conditions were as follows: hot-start at 95 °C for 3 min, denaturation at 95 °C for 1 minute, annealing at 60-65 °C for 1 minute, and extension at 68 °C for 1 minute per 1 kb of amplified cDNA. At the end of 30 cycles, a final extension step was carried out at 68 °C for 8 minutes. Clone J was digested with EcoRI and re-cloned into pcDNA3.1(+) (Invitrogen) and pXJ41neo (Gift from C. Pallen, IMCB, 20) and correctly oriented clones were selected.

Table 1 - Sequences of primers used for PCR amplification of overlapping cDNA regions of the genome of HCV isolate HCV-S1.

Primer	Sequence (5'-3')	Position	Sense/Anti-sense	Reference
H1	ACTGTCTTCACGCAGAAAGCGTCTAGCCAT (SEQ ID NO:1)	-285 to -256	sense	Bukh et al. 1992
H2	CACGCAGAAAGCGTCTAGCCAT (SEQ ID NO:2)	-276 to -247	sense	Bukh et al. 1992
H3	CGAGACCTCCCGGGGCACTCGCAAGCACCC (SEQ ID NO:3)	-14 to -43	anti-sense	Bukh et al. 1992
H4	TCCCGGGGCACTCGCAAGCACCCTATCAGG (SEQ ID NO:4)	-21 to -50	anti-sense	Bukh et al. 1992
H5	CTATCAGGCAGTACCACAAGGCCTTTCGCG (SEQ ID NO:5)	-43 to -72	anti-sense	-
H6	CCCGCYAGGACYCCCCAGTGG (SEQ ID NO:6)	1073 to 1053	anti-sense	Dirsel et al. 1994*
H7	GCCCAGTTCCCCACCATGGA (SEQ ID NO:7)	1106 to 1087	anti-sense	Dirsel et al. 1994*
H8	AGGGCAGTCCTGTTGATGTGC (SEQ ID NO:8)	1280 to 1260	anti-sense	Dirsel et al. 1994*
H9	AGGCTATCATTGCAGTTCAGGGC (SEQ ID NO:9)	1298 to 1276	anti-sense	-
H10	CCACTGGGGRGTCCTRGCGGG (SEQ ID NO:10)	1053 to 1073	sense	Dirsel et al. 1994*
H11	TCCATGGTGGGGAACTGGGC (SEQ ID NO:11)	1087 to 1106	sense	Dirsel et al. 1994*
H12	CGCCTCCGCACGATGCAGCCAT (SEQ ID NO:12)	2461 to 2440	anti-sense	-
H13	AGTACCAGACCTATGAAAACCGC (SEQ ID NO:13)	2483 to 2461	anti-sense	-

H14	ATGGACCGGGAGATGGCTGCA (SEQ ID NO:14)	2428 to 2448	sense	-
H15	AGGCTTTAGCCGTGTGAGACA (SEQ ID NO:15)	4848 to 4828	anti-sense	-
H16	GCGCCYATCACGGCCTACTCC (SEQ ID NO:16)	3079 to 3099	sense	-
H17	GACGACCTCCAGGTCAGCCGA (SEQ ID NO:17)	4968 to 4948	anti-sense	-
H18	ACGCCCACTTCTTGTCTCAGA (SEQ ID NO:18)	4706 to 4726	sense	-
H19	ACTAAGCAGGCAGGAGACAAC (SEQ ID NO:19)	4726 to 4746	sense	-
H20	TTGATGGGTAATTTGCTCTCC (SEQ ID NO:20)	7328 to 7308	anti-sense	-
H21	GTGGTGACGCAGCAAGGAGTT (SEQ ID NO:21)	7359 to 7339	anti-sense	-
H22	CAGCGACGGGTCTTGGTCTAC (SEQ ID NO:22)	7200 to 7220	sense	-
H23	TCACCGGTTGGGGAGCAGATAG (SEQ ID NO:23)	9033 to 9012	anti-sense	-
H24	TCTACGGGGCCTACTACTCCATT (SEQ ID NO:24)	8597 to 8619	sense	-
H25	CTACTACTCCATTGAGCCACTTGAC (SEQ ID NO:25)	8607 to 8631	sense	-
H26	ACATGATCTGCAGAGAGGCCAGTATCA GCACTCTC (SEQ ID NO:26)	9269 to 9234	anti-sense	Tanaka et al. 1996
H27	GTCAAGTGGCTCAATGGAGTAGTAGGC (SEQ ID NO:27)	8631 to 8605	anti-sense	-
H28	GCCAGCCCCCGATTGGGGGCGACACTC CACCATAGATCACTCCCCTGTGAGGAA CTACTGTCTTCACGCAGAAAGCGTCTA GCCA (SEQ ID NO:28)	-341 to -256	sense	-
H29	GACCACGCGTATCGATGTCGACTTTTTT TTTTTTTTTV (SEQ ID NO:29)	-	sense	5'/3' RACE kit
H30	GACCACGCGTATCGATGTCGAC (SEQ ID NO:30)	-	sense	5'/3' RACE kit
H31	ATGGACCAGGAGTTGGCTGCATCGTGC (SEQ ID NO:31)	2769 to-2796	sense	-
H32	CTAACGCGCACGCACGAATGAGGCCTT (SEQ ID NO:32)	3115 to 3008	anti-sense	-

Construction of HCV-S1 cDNA clones encoding the NS proteins:

The region spanning NS3 to NS5A (nucleotides 3420-7669 of SEQ ID NO:33) was obtained by double-cloning a 1.844 kb BamHI/BmrI fragment (nucleotides 3420-5263 of SEQ ID NO:33) from clone F (Figure

1A) and a 2.4 kb BmrI/EcoRV fragment (nucleotides 5263-7669 of SEQ ID NO:33) from clone G (Figure 1A) into pKSII (+/-) digested with BamHI and EcoRV. The resulting clone was digested with XbaI and BsrGI and ligated to a 0.9 kb XbaI/BsrGI fragment (nucleotides 2769-3640 of SEQ ID NO:33) containing the NS2 ORF from clone E, to thereby produce clone K (Figure 1A). To generate the region spanning nucleotides 7200 to 9268 of the HCV genome, clones H and I (Figure 1A) were used as templates in a PCR reaction with primers H22 and H26 (Table 1). The resultant PCR product (nucleotides 7641-9609 of SEQ ID NO:33) was cloned into pCRIITOPO to generate clone L (Figure 1A). Clones K and L were each introduced into electro-competent GM109 bacteria cells and DNA plasmids preparations of these clones were digested with BclI and EcoRV and co-ligated to generate clone M (Figure 1A). Clone M was digested with NotI and XhoI and re-cloned into pcDNA3.1(+) and pXJ41neo to generate pcDNA3(NSP) and pXJ41(NSP) respectively.

Construction of full-length cDNA clones of HCV-S1:

Clones J and M were digested with CspI and XbaI and the resulting 3.3 kb fragment from clone J (nucleotides 1-3369 of SEQ ID NO:33) including the anchor-5'NCR to NS2 sequence was ligated into clone M to generate a full length genome of HCV-S1 in pKSII(+/-) (designated pKSII(S1)). To generate the full length clone in pcDNA3.1(+), the EcoRV/BsrGI fragment from pKSII(S1) was ligated to the pcDNA3(NSP) digested with the same enzymes to generate pcDNA3(S1). The same fragment was cloned into the blunt-NotI/BsrgI site in pXJ41(NSP) to generate pXJ41(S1).

Renilla luciferase expression construct:

The renilla luciferase cDNA (GeneBank Accession number M63501, nucleotide coordinates 10-945) including the upstream intron sequence from human growth hormone (GeneBank Accession number M13438, nucleotide coordinates 569-827) was PCR amplified from pBIND

(Promega) and subcloned into the HindIII site of pcDNA3.1(+). Clones containing the insert in the right orientation were isolated and verified by sequence analysis.

Chimeric HCV-luciferase constructs:

The firefly luciferase gene (GeneBank Accession number M15077, nucleotide coordinates 253-2387) was PCR amplified from the plasmid pGL3-Basic (Promega, Madison, WI). The PCR product was digested with EcoRI and EcoRV and re-cloned into pcDNA3.1(+) (Clontech) to generate the construct pLUCEE(15). The HCV sequence from nt 1- 374 comprising the full length 5'NCR and the first 33 nt of its core sequence (nucleotides 1-374 of SEQ ID NO:33) was PCR amplified from HCV-S1. The PCR product was digested with HindIII and EcoRI and cloned into pLUCEE15 to generate the construct pLUCEE15NC(B2). In order to clone the entire 3'UTR of HCV-S1 downstream of pLUCEE15NC(B2), the plasmid pHCV700(A8) (clone I, Figure 1A) was digested with XcmI and EcoRV and blunted with Klenow. The resultant insert was cloned into the EcoRV site of pLUCEE15NC(B2) and clones with the 3'UTR cloned in the right orientation were isolated. One of these clones pLUCNC3UTR(B9) was excised with HindIII and XhoI, blunted with Klenow and cloned into the EcoRV site of pcDNA3.1(+). Clones with inserts in the anti-sense orientation were isolated and designated pAS9 (Figure 1A). Next, chimeric HCV-luciferase constructs which contained HCV NS5B and 3'UTR sequences were generated. A region covering the C-terminal end of the NS5B sequence and the complete 3'UTR of HCV-S1 was PCR amplified from pHCV700(A8) (clone I, Figure A). The PCR product (nucleotides 9159-9609 of SEQ ID NO:33) was digested with EcoRV and XhoI and cloned into pLUCEE15NC(B2) to generate pLUCNC5BUTR(11). The insert from this construct was excised with HindIII and XhoI, blunted with Klenow and cloned into the EcoRV site of pcDNA3.1(+). Clones with inserts in the anti-sense orientation were isolated and named pAS11 (Figure

1A). All constructs were verified via enzymatic restriction digestions and sequence analyses. Figure 1B illustrates the above described steps utilized in generating the chimeric anti-sense expression constructs pAS9 and pAS11 and their sense oriented counterparts.

Sequence analysis:

DNA sequencing of all constructs was carried out using the Taq DyeDeoxy terminator cycle sequencing kit and an automated DNA sequencer 373 from PE Applied Biosystems (Foster City, CA, USA).

Cells and cell culture:

The human embryonic kidney cell line, 293, its derivative, 293T, which bears the large T antigen from SV40, and the human hepatoma cell line HuH-7 were all purchased from American Type Cell Collection (ATCC). The cells were cultured in Dulbecco's Minimal Essential Media (DMEM) containing 2 mM L-glutamine, and 10% fetal bovine serum and maintained at 37 °C in 5 % CO_2.

Cell transfections:

Transfections were performed using the Effectene™ transfection reagent from QIAGEN (Valencia, CA, USA). Approximately 2×10^5 cells were plated into 6-well tissue culture plates 14-18 hours prior to transfection. A total of 1 µg of plasmid DNA in 150 µl EC bufffer was mixed with 8 µl of enhancer and vortexed for 10 seconds. The mixture was allowed to stand at RT for 2-5 minutes, 25 µl of Effectene™ transfection reagent was added, the mixture vortexed again and incubated at RT for another 5-10 minutes. Cells were washed with PBS, added into DNA-Effectene™ mixture diluted in 2 ml of complete growth medium and incubated at 37°C and 5% CO_2 for 6-8 hours. Following incubation, the medium was removed and the cells were washed with once with PBS. Approximately 2.5 ml of fresh complete medium was added to the cells and the cells were incubated for an aditional 48-120 hours, following which

cells were harvested for RNA isolation or western analysis, or treated with 1000 mg/ml G418 for selection of stable clones.

Luciferase assays:

Luciferase activity was measured using the a luciferase assay kit (Promega, Madison, WI). Following a 72-120 hour incubation period, cells were washed twice with PBS and lysed with 100 µl reporter lysis buffer (Promega). The lysate was allowed to stand at room temperature for 10-15 minutes. Following which, the lysate was centrifuged for 1 min in a microfuge and a 10 µl aliquot was mixed with 100 µl of reporter buffer (Promega); luciferase activity was measured in a Turner luminometer (Turner Designs, Sunnydale, CA) over an integration period of 15 seconds. In cells co-transfected with pCMV-Ren, cell pellets were re-suspended in 100 ml of passive lysis buffer and measured using the dual-luciferase system from Promega. Values obtained were normalized with the levels of Renilla luciferase activity in the cell lysates and the total protein concentration.

***In-vitro* translation:**

Translation was effected via the TNT quick coupled transcription/translation system from Promega. Briefly, 0.5-1 µg of plasmid DNA was mixed with 40 µl of TNT quick master mix and 2 µl of ^{35}S methionine (10mCi/ml) (NEN). The reaction mixture was incubated at 30 °C for 1-3 hours. Following a predetermined time period, an aliquot was removed and SDS-Page analysis was performed. Where indicated, between 0.3-2.5 µl of canine pancreatic microsomal membranes (Promega) were added to the reaction mixture.

Western blot analysis:

Cell lysates were resolved on a 10 or 12% sodium dodecyl sulphate (SDS)-polyacrylamide gel, transferred to a nitrocellulose membrane, blocked with 5% nonfat skim milk in PBS, and incubated with a primary antibody followed by incubation with anti-mouse or anti-human secondary

antibody conjugated to horseradish peroxidase (Sigma). Detection was effected using the ECL enhanced chemiluminescence kit (Pierce). The E2 directed antibody (H52), was a kind gift from J. Dubuisson (Institut de Biologie de Lille & Institut Pasteur de Lille, Lille Cedex, France). The NS3 and NS5A directed monoclonal antibodies were purchased from Devaron, Inc. (NJ, USA) and Biodesign International (ME, USA) respectively.

EXPERIMENTAL RESULTS

Generation of HCV overlapping cDNA clones:

Sera derived from a single chronic HCV carrier were subjected to RT-PCR, and nine overlapping cDNAs clones covering the entire HCV genome were (Figure 1A). The overlapping regions in these clones had almost identical sequences (data not shown). To obtain the complete 5' NCR sequence of this isolate, 5' rapid amplification of cDNA ends was effected using the 5'/3' RACE kit from Boehringer Mannheim. Following two rounds of nested PCR, a cDNA fragment comprising the 5' NCR region spanning nucleotides –341 to –72 that was missing from clones B and C was obtained. The overlapping cDNA clones of isolate HCV-S1 span 9609 nucleotides encoding a complete polyprotein 3010 amino acids long (SEQ ID NO:34), and a 341-nt 5' NCR, and a 235-nt 3' NCR (Figure 1A). To determine the genotype of isolate HCV-S1, the sequence of a region of 226 nt within the 5' NCR (from -276 to -21, Figure 1A) (2) as well as 233 nt within NS3 (from 4699 to 4932, Figure 1A) and 400 nt within NS5B (from 7904 to 8304, Figure 1A) (5) were analyzed. Following comparison to available HCV sequences, it was determined that HCV-S1 belongs to the type 1 genotype, with a 1b subtype. Sequence comparisons of the other two regions were consistent with this finding.

Characterization of full length HCV genome:

The full length HCV genome was generated as described hereinabove to produce pcDNA3(S1) and pXJ41(S1) respectively. To characterize this clone, *in vitro* coupled transcription and translation was

first carried out with pcDNA3(SP) and pcDNA3(NSP) using a kit from Promega. A single polyprotein larger than 185kD was observed following one hour of incubation with pcDNA3(NSP) (Figure 2A, lane 2). Prolonged incubation periods gave rise to smaller protein products (Figure 2A, lanes 4-6). Following two hours of incubation, distinct bands corresponding to proteins of approximately 80, 75 and 62 kD in size were also detected (Figure 2A, lane 6). It is believed that these products are the result of the enzymatic activity of the protease moiety of NS3 and as such these bands possibly correspond to NS3-4A (77kD), NS5B (68kD) and NS5A (58kD).

The construct pcDNA3(SP) contains the entire HCV sequence of the core, E1 and E2 proteins, and the first 115 amino acids of NS2 and as such when translated should give rise to a polyprotein of about 82kD. In vitro translation experiments with this construct with addition of either an enhancer or KCl produced a single band corresponding to about 82kD (Figure 2B, lanes 1-6), whilst addition of magnesium acetate failed to produce any band (Figure 2B, lanes 7-9).

The above described was repeated with the pcDNA3(S1) construct. Following a one hour incubation, a broad band larger than 185kD was observed (Figure 2C, lane 1). Following two hours of incubation, several smaller bands were observed of sizes ranging from 65 to 140kD. In addition, two fainter bands of 60kD and 50kD were also detected (Figure 2C, lane 2). The intensity of the bands increased slightly when incubation was allowed to proceed for three hours (Figure 2C, lane 3). This suggests that the HCV polyprotein was proteolytically cleaved *in vitro*, mostly likely by the NS3 protease. Interestingly addition of canine pancreatic microsomal membranes (CPMM) led to disappearance of the two upper bands of about 140 and 100 kD and reduction in intensity of the lower two bands (Figure 2C, lanes 4 and 5). It is likely that these bands represent subfragments of the HCV polypeptide and were post-translationally processed by the microsomal vesicles. pXJ41(S1) was transiently

transfected into 293T cells, and the expression of HCV proteins was examined. Structural (core and E2) and non-structural (NS3, NS5A, NS5B) proteins (Figure 3A-G) were detected using available monoclonal or polyclonal antibodies.

These results indicate that the full length HCV genome cloned while reducing the present invention to practice, is able to direct the expression of the full length polyprotein and is capable of being processed.

Results of transfection of anti-sense chimeric HCV-luciferase construct pASB9:

The 293T and HuH7 cell lines were separately transfected with two different clones of pASB9 (pASB9.1 and pASB9.2), which contain an anti-sense chimera of the firefly luciferase gene downstream of a HCV 5' NCR-core sequence and upstream of the HCV 3' UTR sequence. Transfection was carried out with pASB9 and an equal amount of pXJ41(NSP), pXJ41(S1) or a combination of pXJ41(NS3) and pXJ41(NS5B). Co-transfection with the vector, pXJ41neo was used as a control to measure background luciferase activity. The cells were harvested and assayed for luciferase activity 5 days post-transfection. There was no observed increase in luciferase activity in co-transfection experiments with any of the HCV expression constructs compared to co-transfection with the vector (data not shown). Experiments carried out with a total of 1 or 2 ? g of DNA produced similar results.

Results of transfection of anti-sense chimeric HCV-luciferase construct pAS11:

Similar experiments were carried with the anti-sense construct pAS11 (pAS11-12 and pAS11-15) which contains the anti-sense chimera of the firefly luciferase gene downstream of a HCV 5' NCR-core sequence and upstream of the HCV NS5B-3' UTR sequence. In 293T cells, co-transfection with the full length HCV expression plasmid, pXJ41(S1) and pAS11-12 produced a 10-fold increase over background luciferase

activities five days post transfection, while a 14.7-fold increase was observed with pAS11-15 co-transfected with pXJ41(S1) (Table 2). In similarly transfected HuH7 cells, luciferase activities were 2.7-fold and 5.8-fold above background values three days post transfection (Table 3). At five days post transfection, the luciferase activities in HuH7 cells slightly increased to 3.7-fold and 6.2-fold respectively (Table 4). However, co-transfection of pAS11-12 or -15 with the NS proteins expression vector (pXJ41(NSP)) or the vector including NS3 and NS5B, resulted in no detectable increase in luciferase activity as compared to transfection with vector alone (Tables 2-4).

Table 2

	Constructs	R1	R2	Av	
1	pAS11(12)+pXJ41neo	1.20	1.06	1.13	
2	pAS11(12)+pXJ41(NS3)+pXJ41(NS5B)	1.22	1.46	1.34	
3	pAS11(12)+pXJ41(NSP)	1.40	1.30	1.35	
4	pAS11(12)+pXJ41(S1)	12.34	11.11	11.73	10.36X
5	pAS11(15)+vector	1.39	1.46	1.42	
6	pAS11(15)+pXJ41(NS3)+pXJ41(NS5B)	1.92	1.53	1.73	
7	pAS11(15)+pXJ41(NSP)	1.31	1.33	1.32	
8	pAS11(15)+pXJ41(S1)	21.79	20.09	20.94	14.7X
9	p11(3)+pXJ41neo	5106.2	5082.7	5094.5	
10	p11(6)+pXJ41neo	5611.5	5440.0	5525.8	
11	pXJ41neo	0	-0.2	0	

p11(3) - sense, clone 11

p11(6) - sense, clone p11

R - luciferase reading

Av - average luciferase reading

Table 3

	Constructs	R1	R2	Av	
1	pAS11(12)+pXJ41neo	5.52	5.27	5.40	
2	pAS11(12)+pXJ41(NS3)+pXJ41(NS5B)	3.10	2.98	3.04	

3	pAS11(12)+pXJ41(NSP)	4.11	4.12	4.11	
4	pAS11(12)+pXJ41(S1)	15.21	14.26	14.74	2.73X
5	pAS11(15)+vector	3.25	3.70	3.47	
6	pAS11(15)+pXJ41(NS3)+pXJ41(NS5B)	3.15	3.23	3.19	
7	pAS11(15)+pXJ41(NSP)	3.24	3.14	3.19	
8	pAS11(15)+pXJ41(S1)	19.87	20.07	19.97	5.76X
9	p11(3)+pXJ41neo	10348.8	9848.0	10098.4	
10	p11(6)+pXJ41neo	13418.2	12821.6	13119.9	
11	pXJ41neo	0	0	0	

p11(3) - sense, clone 11

p11(6) - sense, clone p11

R - luciferase reading

Av - average luciferase reading

Table 4

	Constructs	R1	R2	Av	
1	pAS11(12)+pXJ41neo	2.03	2.09	2.06	
2	pAS11(12)+pXJ41(NS3)+pXJ41(NS5B)	0.82	0.66	0.74	
3	pAS11(12)+pXJ41(NSP)	1.36	1.09	1.22	
4	pAS11(12)+pXJ41(S1)	8.26	6.98	7.62	3.7X
5	pAS11(15)+vector	2.19	2.02	2.10	
6	pAS11(15)+pXJ41(NS3)+pXJ41(NS5B)	0.92	0.79	0.86	
7	pAS11(15)+pXJ41(NSP)	0.63	0.61	0.62	
8	pAS11(15)+pXJ41(S1)	13.89	12.28	13.09	6.23X
9	p11(3)+pXJ41neo	3565.8	3450.3	3508.1	
10	p11(6)+pXJ41neo	4197.2	3861.9	4030	
11	pXJ41neo	0	0	0	

p11(3) - sense, clone 11

p11(6) - sense, clone p11

R - luciferase reading

Av - average luciferase reading

Results of co-transfection with pAS11 and pCMV-Ren:

Similar experiments were conducted using a renilla expression construct pCMV-Ren, in order to account for any variation in luciferase activity due to different transfection efficiencies. 293T and HuH7 cells were transfected with a total of 1 μg of DNA and cells were harvested and analyzed 3 days post-transfection. All values obtained were normalized against total protein concentration and renilla luciferase activity. In 293T cells, co-transfection of pAS11-15 with pXJ41(S1) resulted in a 18.5-fold increase over background luciferase activity (Table 5). In HuH7 cells, the luciferase activity of pAS11-15 was 3.9-fold higher when co-transfected with pXJ41(S1) (Table 6).

Table 5

		5X dilution	5X dilution		5X dilution	Neat	Neat
	Constructs	Av FF LUC	Av REN	N. Ren (X)	N. Av FFL	Av FFL	Total FFL
1	pAS11(12)+pXJ41neo	0.39	1491	1.00	0.39	8.56	8.56
2	pAS11(12)+pXJ41(S1)	0.72	478	3.12	2.25	145.80	158.146
3	AS11(3)+pXJ41neo	37.13	1626	0.92	34.16	9538.00	11445.6
	Constructs	Prot. Conc	PN (X)	Total FFL	Final FFL		
1	pAS11(12)+pXJ41neo	0.303	1	8.56	8.56		
2	pAS11(12)+pXJ41(S1)	0.284	1.07	145.80	158.15	18.48X	
3	AS11(3)+pXJ41neo	0.253	1.2	9538.00	11445.60		

Av FF LUC = average of 2 firefly luciferase readings from 20 ml of 5X diluted cell lysate

Av REN = average of 2 renilla luciferase readings from 20 ml of 5X diluted cell lysate

N. Ren (X) = normalisation index of renilla luciferase readings from 20 ml of 5X diluted cell lysate

N. Av FFL= average of 2 firefly luciferase readings from 20 ml of 5X diluted cell lysate after normalisation against renilla luciferase index

Av FFL= normalised average of 2 firefly luciferase readings from 20 ml of cell lysate (neat)

Final Total FFL= normalised average of 2 firefly luciferase readings of total cell lysate

PN (X)= protein normalisation index

Final FFL= final average of 2 firefly luciferase readings of total cell lysate

Table 6

		5X dilution	5X dilution		5X dilution	Neat	Neat
	Constructs	Av FF LUC	Av REN	N. Ren (X)	N. Av FFL	Av FFL	Total FFL
1	pAS11(12)+pXJ41neo	0.39	1491	1.00	0.39	1.97	7.86
2	pAS11(12)+pXJ41(S1)	0.72	478	3.12	2.25	11.23	44.92
3	AS11(3)+pXJ41neo	37.13	1626	0.92	34.16	17.08	683.20
	Constructs	Prot. Conc	PN (X)	Total FFL	Final FFL		
1	pAS11(12)+pXJ41neo	0.063	1	7.86	7.86		
2	pAS11(12)+pXJ41(S1)	0.093	0.68	44.92	30.55	3.88X	
3	AS11(3)+pXJ41neo	0.106	0.59	683.20	403.09		

Av FF LUC = average of 2 firefly luciferase readings from 20 ml of 5X diluted cell lysate

Av REN = average of 2 renilla luciferase readings from 20 ml of 5X diluted cell lysate

N. Ren (X) = normalisation index of renilla luciferase readings from 20 ml of 5X diluted cell lysate

N. Av FFL= average of 2 firefly luciferase readings from 20 ml of 5X diluted cell lysate after normalisation against renilla luciferase index

Av FFL= normalised average of 2 firefly luciferase readings from 20 ml of cell lysate (neat)

Final Total FFL= normalised average of 2 firefly luciferase readings of total cell lysate

PN (X)= protein normalisation index

Final FFL= final average of 2 firefly luciferase readings of total cell lysate

Cells co-transfected with pASB9 with different HCV expression constructs failed to produce changes in luciferase activity (data not shown). However, pAS11 consistently produced increased luciferase activities when co-transfected with pXJ41(S1), which expresses the full length HCV genome. In 293T cells, the levels were between 10.4-14.7 folds above background levels, and in HuH7 cells they were between 2.7-6.2 folds (Tables 2-4). Even after normalizing with co-transfection with a plasmid that expresses renilla luciferase, a significant increase in luciferase activities was observed. In 293T cells, the increase was 17-fold above background, while in HuH7 cells, it was 3.9-fold (Tables 5 and 6). These results indicate that the additional C-terminal NS5B coding sequence present only in

pAS11 is important and necessary for the NS5B polymerase (and perhaps other factors) to bind efficiently and initiate reverse strand synthesis.

Several reports have shown that *in vitro* provided NS5B is capable of binding and initiating the synthesis of sequences containing the 3' UTR alone (17, 18). Yet, the experiments conducted while reducing the present invention to practice clearly indicate that the 3' UTR alone is insufficient in promoting polymerase activity *in vivo*. As such, this is the first demonstration that the NS5B region works together with the 3' UTR to facilitate negative strand synthesis *in vivo*.

Interestingly co-transfection with an expression vector for the non-structural proteins, pXJ41(NSP) or with expression vectors for NS3 and NS5B did not result in any increase in luciferase activity when compared to co-transfection with the vector alone. This suggests that the synthesis of the sense strand HCV-luciferase chimeric RNA by the HCV NS5B polymerase is dependent on multiple viral proteins, including both non-structural and viral protein(s). It also indicates that a full length replication-competent HCV genome is required for this assay to be functional.

This is the first demonstration that negative strand synthesis depends on expression of essentially all the viral proteins in intact cells. Based on these findings, the present invention provides a cell-based HCV replication-dependent system that is a measure of the activity of the full-length HCV genome. This system is simple, and robust and highly reproducible and in addition, enables to measure viral activity as early as three days post-transfection.

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications

and variations that fall within the spirit and broad scope of the appended claims. All publications, patents, patent applications and sequences disclosed therein and/or identified by a GeneBank accession number mentioned in this specification are herein incorporated in their entirety by reference into the specification, to the same extent as if each individual publication, patent, patent application or sequence was specifically and individually indicated to be incorporated herein by reference. In addition, citation or identification of any reference in this application shall not be construed as an admission that such reference is available as prior art to the present invention.

REFERENCES

1. Houghton M, Weiner A, Han J, Kuo. G and Choo QL 1991 Hepatology 14: 381-388.

2. Alter HJ, Purcell RH, Shih JW, Melpolder JC, Houghton M, Choo QL, Kuo G (1989) Detection of antibody to hepatitis C virus in prospectively followed transfusion receipients with acute and chronic non-A, non-B hepatitis. N Engl. J Med. 321: 1491-1500.

3. Choo QL, Richman KH, Han JH, Berger K, Lee C, Dong C, Gallegos C, Coit D, Medina SR, Barr PJ (1991) Genetic organization and diversity of the hepatitis virus. PNAS 88: 2451-2455.

4. (4) Harada S, Watanabe Y, Takeuchi K, Suzuki T, Katayama T, Takebe Y, Saito et al. 1991 J Virol. 65: 3015-3021.

5. (5) Hijikata M, Kato N, Ootsuyama Y, Nakagawa M, Shimotohno K 1991. Gene mapping of the putative structural region of the hepatitis C virus genome. PNAS 88: 5547-5555.

6. Matsura Y, Harada S, Suzuki R, Watanabe Y, Inoue Y, Saito , Miyamura T 1992. Expression of processed envelope protein hepatitis C virus in mammalian and insect cells. J. Virol. 66: 1425-1431.

7. Tomei L, Failla C, Santolini E, De-Francesco R, La-Monica N 1993 NS3 is a serine protease required for processing of hepatitis C virus polyprotein. J. Virol. 67: 4017-4026.

8. Selby Mj, Choo QL, Berger K. Kuo G, Glazer E., Eckart M, Lee C et al. 1993 Expression, identification and subcellular localisation of the proteins encoded by the hepatitis C viral genome. J. Gen Virol. 74: 1103-1113.

9. Shimizu YK, Purcell RH, Yoshikura H. 1993. Correlation between the infectivity of hepatitis C virus in vivo and its infectivity in vitro. PNAS 90: 6037-6041.

10. Bertolini L, Iacovacci S, Ponzetto A, Gorini G, Bataglia M, Carloni G. 1993 The humna bone marrow-derived B-cell line susceptible to hepatitis C virus infection. Res. Virol. 144: 281-285.

11. Iacovacci S, Sargiacomo M, Parolini I, Ponzetto A, Peschle C, Carloni G. 1993 Replication and multiplication of hepatitis C virus genome in human fetal liver cells. Res Virol. 144: 275-279.

12. Ito T, Mukaigawa J, Zuo J, Hirabayashi Y, Mitamura K, Yasui K. 1996 Cultivation of hepatitis C virus in primary hepatocyte culture from patients with chronic hepatitis C results in release of high titer infectious virus. J. Gen Virol. 77: 1043-1054.

13. Lanford RE, Sureau C, Jacob JR, White R, Fuerst TR. 1994. Demonstration of in vitro infection of hepatocytes with hepatitis C virus using starnd-specific RT/PCR. Virology 202: 606-614.

14. Lohmann, V, Korner F, Koch JO, Herian U, Theilmann L, Bartenschlager R. 1999 Replication of subgenomic Hepatitis C virus RNAs in a hepatoma cell line. Science 285: 110-113.

15. Behrens SE, Tomei L, De Francesco R. 1996. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. EMBO J 2;15(1):12-22.

16. Sun XL, Johnson RB, Hockman MA, Wang QM. 2000. De novo RNA synthesis catalyzed by HCV RNA-dependent RNA polymerase. Biochem Biophys Res Commun. 268(3):798-803.

17. Zhong W, Uss AS, Ferrari E, Lau JY, and Z. Hong. 2000. De novo initiation of RNA synthesis by hepatitis C virus nonstructural protein 5B polymerase. J Virol 74(4):2017-22.

18. Oh JW, Sheu GT, and MM. Lai. 2000. Template requirement and initiation site selection by hepatitis C virus polymerase on a minimal viral RNA template. J Biol Chem. 2000 Apr 3.

19. Zheng XM, Wang Y, Pallen CJ. 1992 Cell transformation and activation of pp60c-src by overexpression of a protein tyrosine phosphatase. Nature. 359: 336-9.

WHAT IS CLAIMED IS:

1. A nucleic acid construct comprising:

(a) an expression cassette including:

(i) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of said RNA virus;

(ii) a second polynucleotide region including a 3' UTR sequence of said RNA virus and at least a C-terminal portion of a coding sequence of said virus; and

(iii) a third polynucleotide region encoding a reporter molecule, said third polynucleotide region being flanked by said first and said second polynucleotide regions; and

(b) a promoter sequence being operatively linked to said expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from said expression cassette.

2. The nucleic acid construct of claim 1, wherein at least a portion of said first polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 1-374 of SEQ ID NO:33.

3. The nucleic acid construct of claim 1, wherein at least a portion of said second polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 9158-9609 of SEQ ID NO:33.

4. The nucleic acid construct of claim 1, wherein said first polynucleotide region further includes a 5' UTR sequence of said RNA virus.

5. The nucleic acid construct of claim 1, wherein said C-terminal portion of said coding sequence of said virus includes coding sequences of a polymerase of said virus.

6. The nucleic acid construct of claim 1, wherein said first polynucleotide region includes an IRES sequence.

7. The nucleic acid construct of claim 1, wherein said RNA virus is selected from the group consisting of a positive strand RNA virus and a negative strand RNA virus.

8. The nucleic acid construct of claim 1, wherein said RNA virus is selected from the group consisting of a virus of the picornavirus family, a virus of the togavirus family, a virus of the orthomyxovirus family, a virus of the paramyxovirus family, a virus of the coronavirus family, a virus of the calicivirus family, a virus of the arenavirus family, a virus of the rhabdovirus family and a virus of the bunyavirus family.

9. The nucleic acid construct of claim 1, wherein said RNA virus is Hepatitis C.

10. The nucleic acid construct of claim 1, wherein said first and said second polynucleotide regions are selected such that said minus strand RNA molecule transcribable from said expression cassette is replicatable by an RNA dependent RNA polymerase of said virus into a plus strand RNA molecule.

11. The nucleic acid construct of claim 1, wherein said promoter is functional in a eukaryotic cell.

12. The nucleic acid construct of claim 11, wherein said eukaryotic cell is selected from the group consisting of an insect cell, a yeast cell and a mammalian cell.

13. The nucleic acid construct of claim 1, wherein said reporter molecule is a polypeptide selected from the group consisting of an enzyme, a fluorophore, a substrate and a ligand.

14. A genetically transformed cell comprising a nucleic acid construct including:

(a) an expression cassette including:

(i) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of said RNA virus;

(ii) a second polynucleotide region including a 3' UTR sequence of said RNA virus and at least a C-terminal portion of a coding sequence of said virus; and

(iii) a third polynucleotide region encoding a reporter molecule, said third polynucleotide region being flanked by said first and said second polynucleotide regions; and

(b) a promoter sequence being operatively linked to said expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from said expression cassette.

15. The genetically transformed cell of claim 14, further comprising an additional nucleic acid construct for expressing at least an RNA dependent RNA polymerase of a virus, said first and said second polynucleotide regions being selected such that said RNA dependent RNA

polymerase is capable of replicating said minus strand RNA molecule into plus strand RNA.

16. The genetically transformed cell of claim 14, wherein at least a portion of said first polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 1-374 of SEQ ID NO:33.

17. The genetically transformed cell of claim 14, wherein at least a portion of said second polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 9158-9609 of SEQ ID NO:33.

18. A method of detecting a presence of an RNA virus in a cell, the method comprising the steps of:

(a) incubating a nucleic acid construct with an extract of the cell under conditions suitable for transcription and translation of said nucleic acid construct, said nucleic acid construct including:

(i) an expression cassette having:

(one) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of said RNA virus;

(two) a second polynucleotide region including a 3' UTR sequence of said RNA virus and at least a C-terminal portion of a coding sequence of said virus; and

(three) a third polynucleotide region encoding a reporter molecule, said third polynucleotide region being flanked by said first and said second polynucleotide regions; and

(ii) a promoter sequence being operatively linked to said expression cassette in a manner so as to direct the transcription of a minus strand RNA molecule from said expression cassette when said nucleic acid construct is incubated with said extract, said first and said second polynucleotide regions being selected such that said minus strand RNA molecule transcribed is replicatable by a polymerase of the RNA virus into a plus strand RNA molecule; and

(b) quantifying a level of said reporter molecule to thereby determine the presence of the virus in the cell.

19. The method of claim 18, wherein said reporter molecule is a polypeptide translated from said plus strand RNA molecule.

20. The method of claim 18, further comprising the step of comparing said level of said reporter molecule to that obtained from cells free of the virus.

21. The method of claim 18, wherein at least a portion of said first polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 1-374 of SEQ ID NO:33.

22. The method of claim 18, wherein at least a portion of said second polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 9158-9609 of SEQ ID NO:33.

23. A method of detecting the presence of an RNA virus in a cell, the method comprising the steps of:

(a) expressing a nucleic acid construct within the cell, said nucleic acid construct including:

(i) an expression cassette having:

(one) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of said RNA virus;

(two) a second polynucleotide region including a 3' UTR sequence of said RNA virus and at least a C-terminal portion of a coding sequence of said virus; and

(three) a third polynucleotide region encoding a reporter molecule, said third polynucleotide region being flanked by said first and said second polynucleotide regions; and

(ii) a promoter sequence being operatively linked to said expression cassette in a manner so as to direct the transcription of a minus strand RNA molecule from said expression cassette when said nucleic acid construct is expressed within the cell, said first and said second polynucleotide regions being selected such that said minus strand RNA molecule transcribed is replicatable by a polymerase of the RNA virus into a plus strand RNA molecule; and

(b) quantifying a level of said reporter molecule to thereby determine the presence of the virus in the cell.

24. The method of claim 23, wherein said reporter molecule is a polypeptide translated from said plus strand RNA molecule.

25. The method of claim 23, further comprising the step of comparing said level of said reporter molecule to that obtained from cells free of the virus.

26. The method of claim 23, wherein at least a portion of said first polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 1-374 of SEQ ID NO:33.

27. The method of claim 23, wherein at least a portion of said second polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 9158-9609 of SEQ ID NO:33.

28. A method of screening for anti-viral drugs, the method comprising the steps of:

(a) co-incubating a nucleic acid construct, a polynucleotide encoding at least a polymerase of an RNA virus and a potential anti-viral molecule under conditions suitable for transcription and translation of said nucleic acid construct and said polynucleotide encoding at least said polymerase, said nucleic acid construct including:

(i) an expression cassette having:

(one) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of said RNA virus;

(two) a second polynucleotide region including a 3' UTR sequence of said RNA virus and at least a C-terminal portion of a coding sequence of said virus; and

(three)a third polynucleotide region encoding a reporter molecule, said third polynucleotide region being flanked by said first and said second polynucleotide regions; and

(ii) a promoter sequence being operatively linked to said expression cassette in a manner so as to direct the transcription of a minus strand RNA molecule from said expression cassette when said nucleic acid construct is incubated with said polynucleotide encoding at least said polymerase of said RNA virus under said conditions suitable for transcription and translation, said first and said second polynucleotide regions being selected such that said minus strand RNA molecule transcribed is replicatable by said polymerase of said RNA virus into a plus strand RNA molecule; and

(b) quantifying a level of said reporter molecule to thereby determine the anti-viral activity of said potential anti-viral molecule.

29. The method of claim 28, wherein said reporter molecule is a polypeptide translated from said plus strand RNA molecule.

30. The method of claim 28, further comprising the step of comparing said level of said reporter molecule to that obtained from cells free of the virus.

31. The method of claim 28, wherein said potential anti-viral molecule is selected from the group consisting of a nucleoside or a nucleotide analogue and an immune-modulatory molecule.

32. The method of claim 28, wherein step (a) is effected by introducing said nucleic acid construct, said polynucleotide encoding at least said polymerase of said RNA virus and said potential anti-viral molecule into a cell.

33. The method of claim 28, wherein step (a) is effected by introducing said nucleic acid construct and said potential anti-viral molecule into a cell infected with said RNA virus.

34. The method of claim 28, wherein at least a portion of said first polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 1-374 of SEQ ID NO:33.

35. The method of claim 28, wherein at least a portion of said second polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 9158-9609 of SEQ ID NO:33.

36. A method of determining drug resistance of an RNA virus, the method comprising the steps of:

(a) co-incubating a nucleic acid construct, a polynucleotide encoding at least a polymerase of the RNA virus and an anti-viral drug molecule under conditions suitable for transcription and translation of said nucleic acid construct and said polynucleotide encoding at least said polymerase, said nucleic acid construct including:

(i) an expression cassette having:

(one) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a coding sequence of said RNA virus;

(two) a second polynucleotide region including a 3' UTR sequence of said RNA virus and at least a C-terminal portion of a coding sequence of said virus; and

(three) a third polynucleotide region encoding a reporter molecule, said third polynucleotide region being flanked by said first and said second polynucleotide regions; and

(ii) a promoter sequence being operatively linked to said expression cassette in a manner so as to direct the transcription of a minus strand RNA molecule from said expression cassette when said nucleic acid construct is incubated with said polynucleotide encoding at least said polymerase of the RNA virus under said conditions suitable for transcription and translation, said first and said second polynucleotide regions being selected such that said minus strand RNA molecule transcribed is replicatable by said polymerase of the RNA virus into a plus strand RNA molecule; and

(b) quantifying a level of said reporter molecule to thereby determine the resistance of the RNA virus to said anti-viral drug.

37. The method of claim 36, further comprising the step of comparing said level of said reporter molecule to that obtained from cells free of said anti-viral drug.

38. The method of claim 36, wherein said reporter molecule is a polypeptide translated from said plus strand RNA molecule.

39. The method of claim 36, wherein said anti-viral drug is selected from the group consisting of a nucleoside or nucleotide analog and an immune-modulatory molecule.

40. The method of claim 36, wherein step (a) is effected by introducing said nucleic acid construct, said polynucleotide encoding at least said polymerase of said RNA virus and said anti-viral drug into a cell.

41. The method of claim 36, wherein step (a) is effected by introducing said nucleic acid construct and said anti-viral drug into a cell infected with the RNA virus.

42. The method of claim 36, wherein at least a portion of said first polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 1-374 of SEQ ID NO:33.

43. The method of claim 36, wherein at least a portion of said second polynucleotide region is at least 50 % identical to a sequence encompassed by nucleotides 9158-9609 of SEQ ID NO:33.

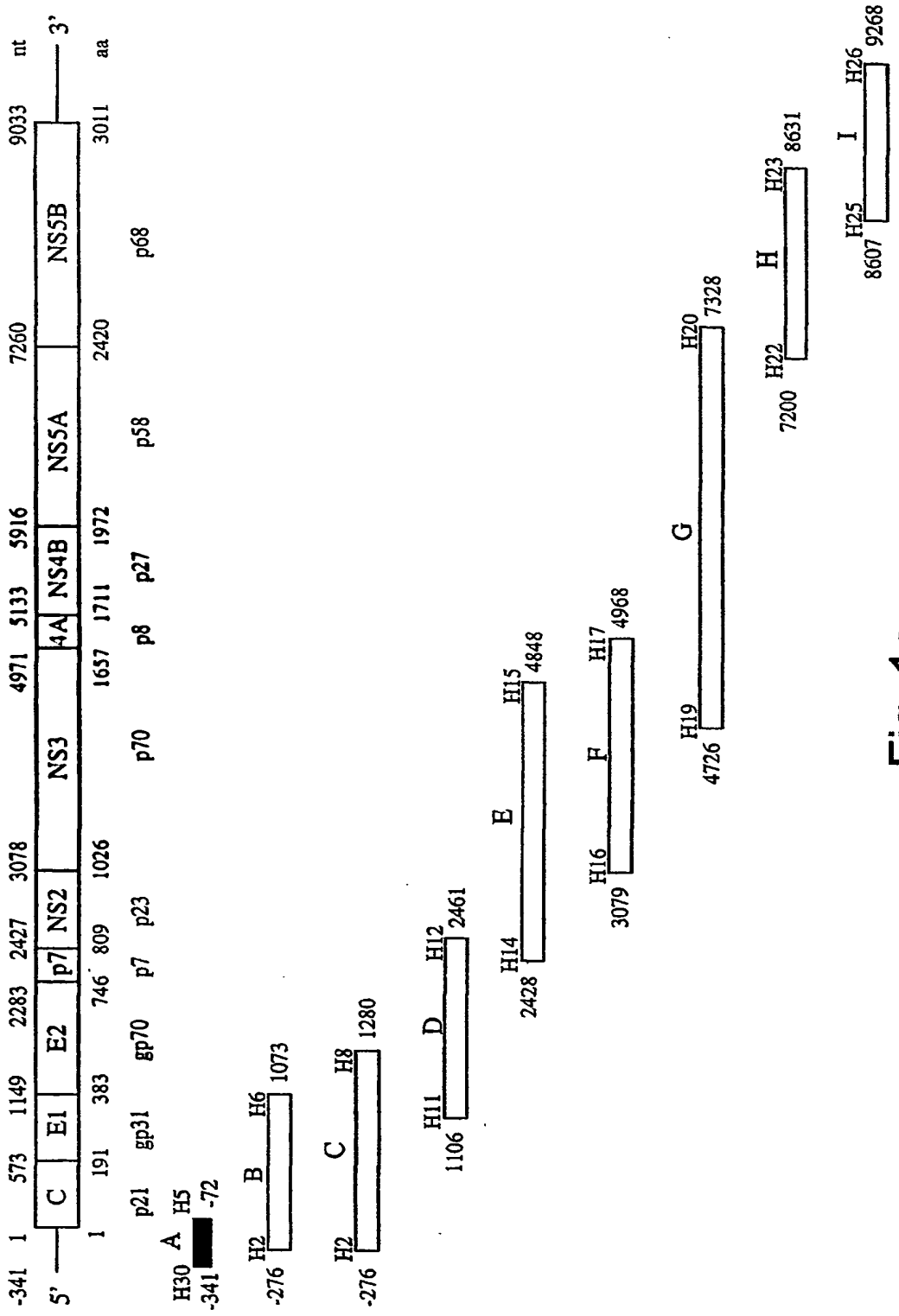

Fig. 1a

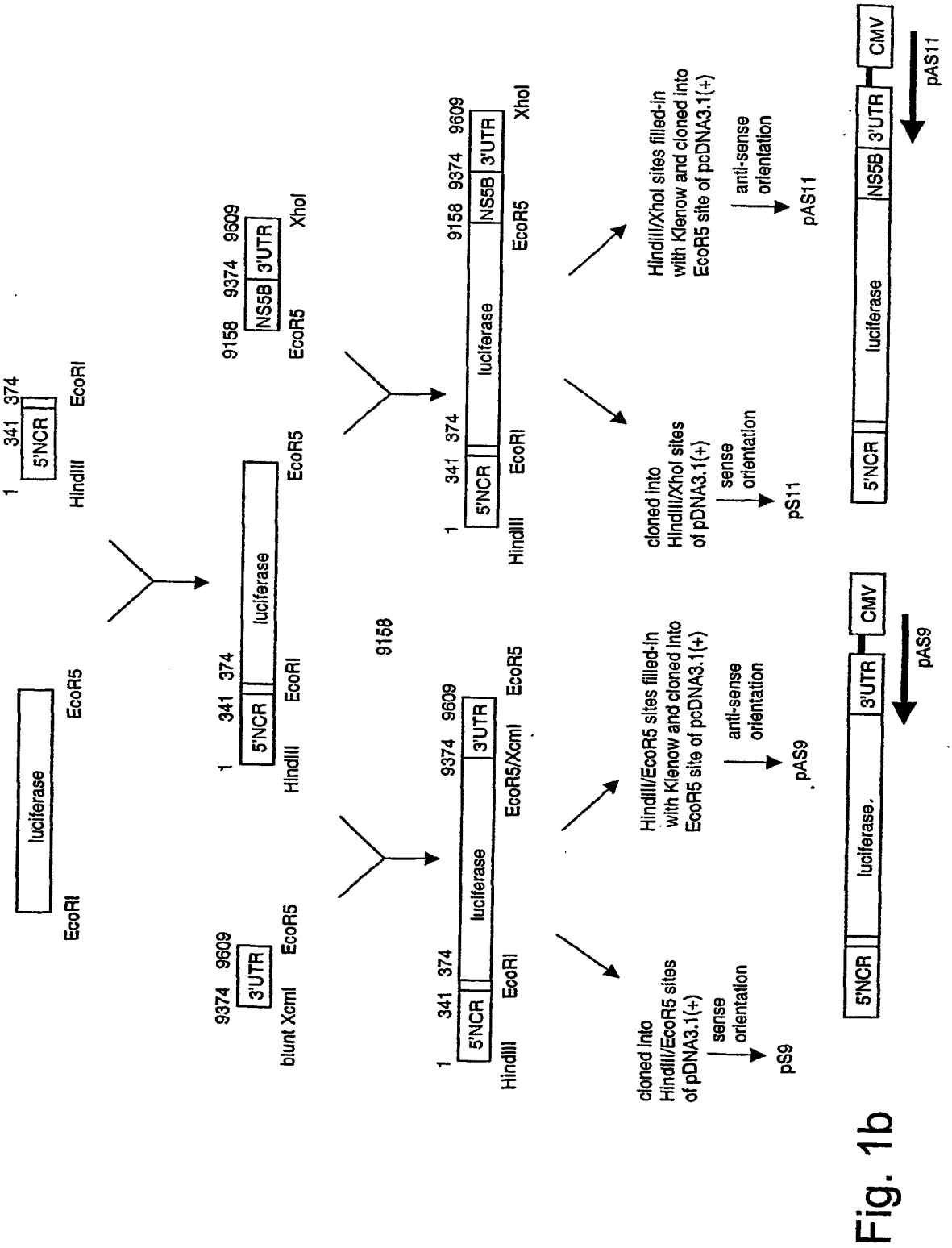

Fig. 1b

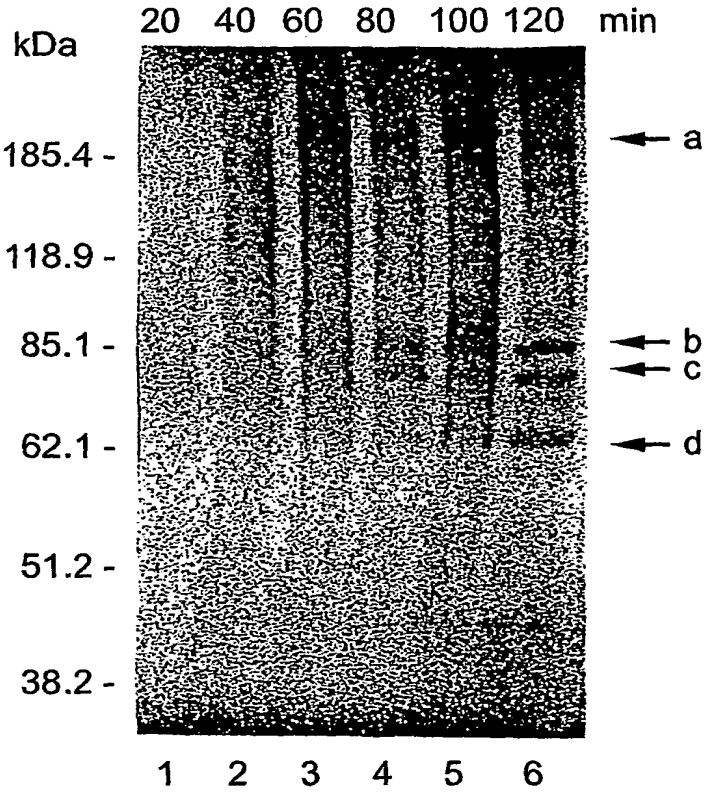

Fig. 2a

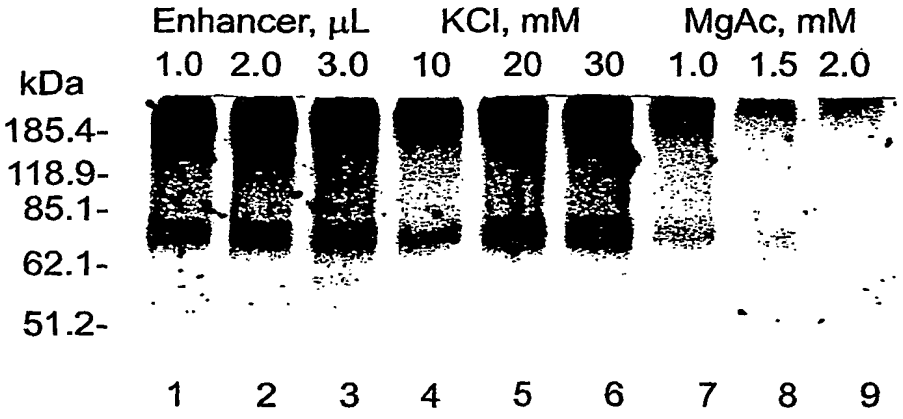

Fig. 2b

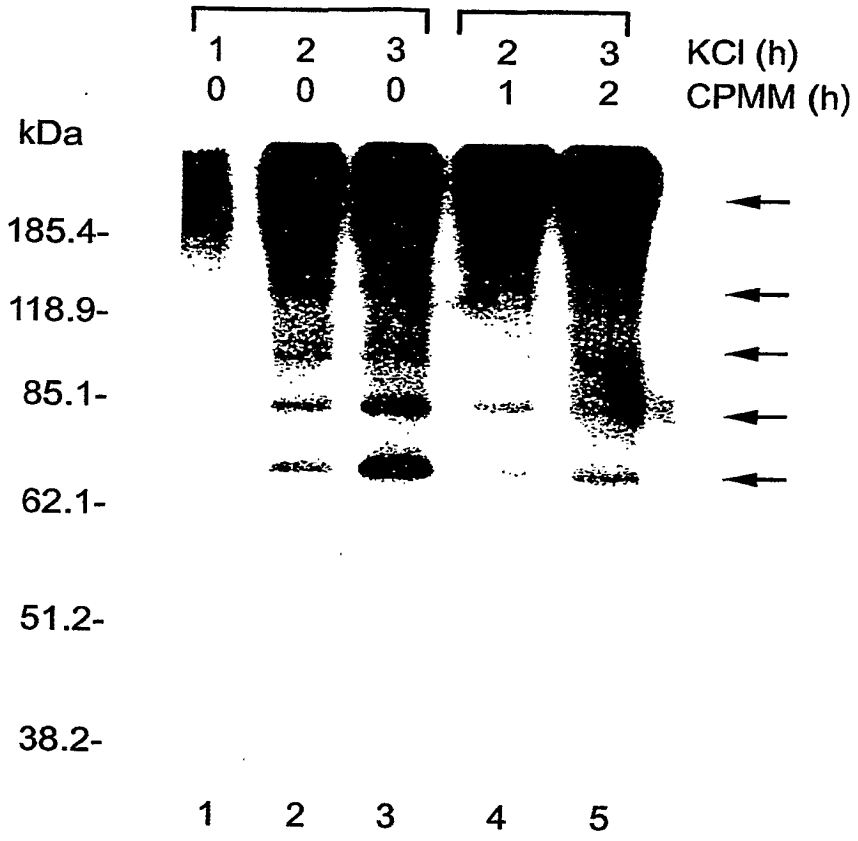

Fig. 2c

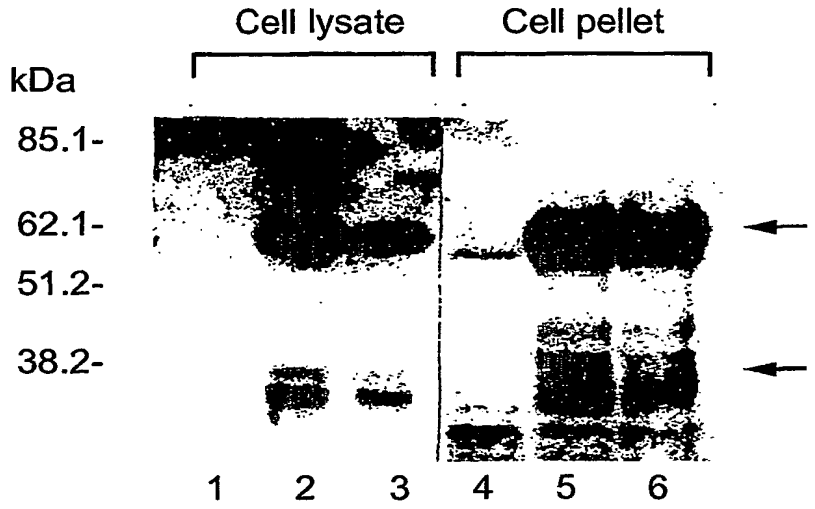

Fig. 3a

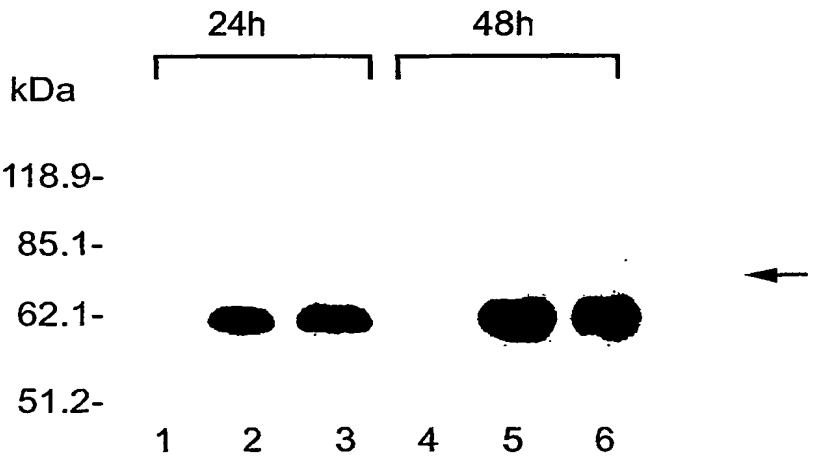

Fig. 3b

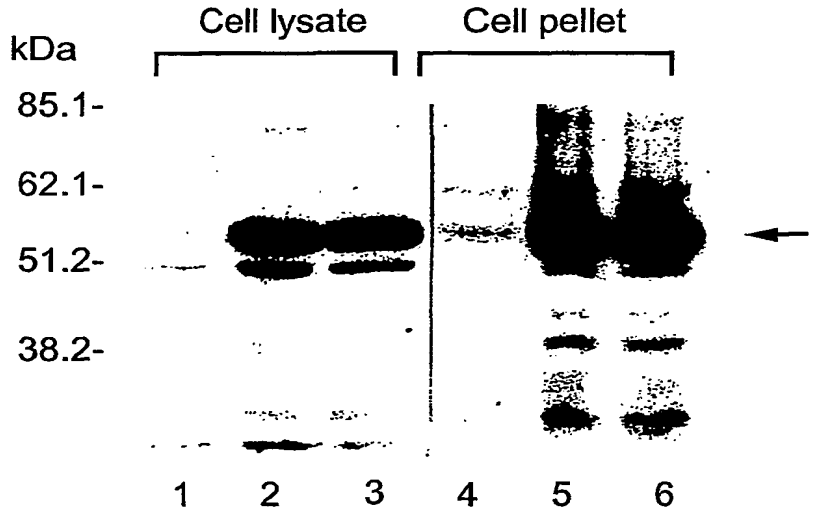

Fig. 3c

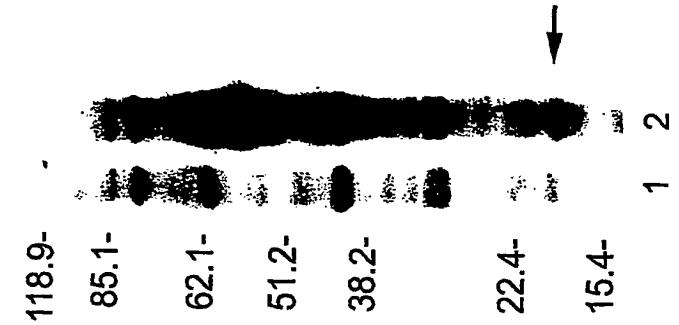

Fig. 3g

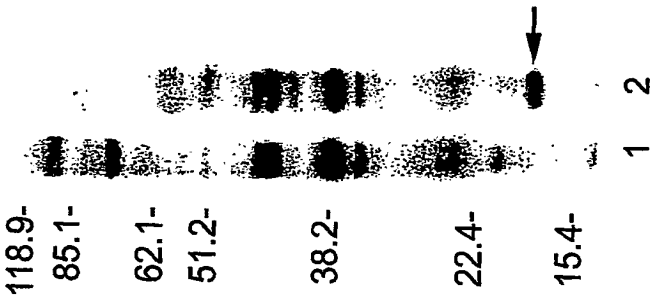

Fig. 3f

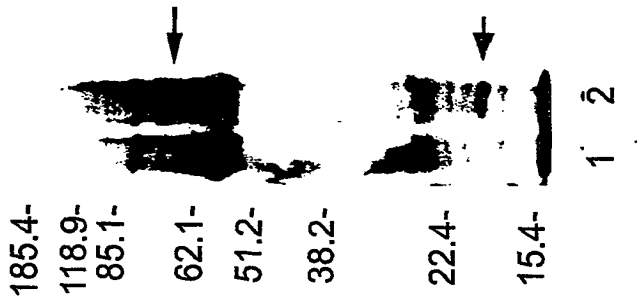

Fig. 3e

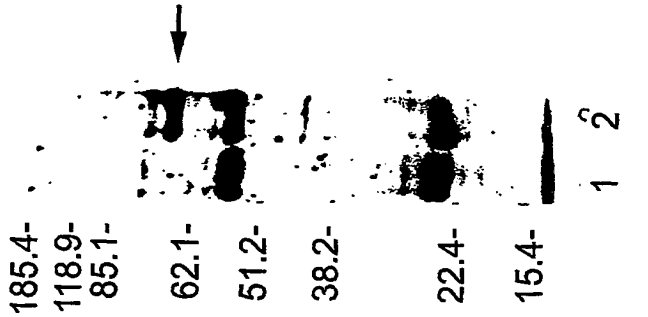

Fig. 3d

SEQUENCE LISTING

<110> Siew Pheng Lim et al.

<120> NUCLEIC ACID CONSTRUCTS AND METHODS OF UTILIZING SAME FOR DETECTING VIRAL INFECTION, FOR UNCOVERING ANTI-VIRAL DRUG CANDIDATES AND FOR DETERMINING DRUG RESISTANCE OF VIRAL ISOLATES

<130> 01/22137

<160> 32

<170> PatentIn Ver. 2.1

<210> 1

<211> 30

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial sequence: PCR primer

<400> 1

ACTGTCTTCA CGCAGAAAGC GTCTAGCCAT 30

<210> 2

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial sequence: PCR primer

<400> 2

CACGCAGAAA GCGTCTAGCC AT 22

<210> 3

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial sequence: PCR primer

<400> 3

CGAGACCTCCC GGGGCACTCG CAAGCACCC 29

<210> 4

<211> xx

<212> DNA

<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 4
TCCCGGGGCACTCGCAAGCACCCTATCAGG

<210> 5
<211> 30
<212> DNA
<213> Artificial sequence
<220>
<223> Description of Artificial sequence: PCR primer
<400> 5
CTATCAGGCA GTACCACAAG GCCTTTCGCG 30

<210> 6
<211> 20
<212> DNA
<213> Artificial sequence
<220>
<223> Description of Artificial sequence: PCR primer
<400> 6
CCCGCYAGGA CYCCCCAGTGG 20

<210> 7
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 7
GCCCAGTTCC CCACCATGGA 20

<210> 8
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 8
AGGGCAGTCC TGTTGATGTG C 21

<210> 9
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 9
AGGCTATCAT TGCAGTTCAG GGC 23

<210> 10
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 10
CCACTGGGGR GTCCTRGCGG G 21

<210> 11
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 11
TCCATGGTGG GGAACTGGGC 20

<210> 12
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 12
CGCCTCCGCA CGATGCAGCC AT 22

<210> 13
<211> 23

<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 13
AGTACCAGAC CTATGAAAAC CGC 23

<210> 14
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 14
ATGGACCGGG AGATGGCTGC A 21

<210> 15
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 15
AGGCTTTAGC CGTGTGAGAC A 21

<210> 16
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 16
GCGCCYATCA CGGCCTACTC C 21

<210> 17
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 17
GACGACCTCC AGGTCAGCCG A 21

<210> 18
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 18
ACGCCCACTT CTTGTCTCAG A 21

<210> 19
<211> xx
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 19
ACTAAGCAGG CAGGAGACAA C 21

<210> 20
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 20
TTGATGGGTA ATTTGCTCTC C 21

<210> 21
<211> xx
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 21
GTGGTGACGC AGCAAGGAGT T 21

<210> 22
<211> xx
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 22
CAGCGACGGG TCTTGGTCTA C 21

<210> 23
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 23
TCACCGGTTG GGGAGCAGAT AG 22

<210> 24
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 24
TCTACGGGGC CTACTACTCC ATT 23

<210> 25
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 25
CTACTACTCC ATTGAGCCAC TTGAC 25

<210> 26
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 26
ACATGATCTG CAGAGAGGCC AGTATCAGCA CTC 33

<210> 27
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 27
GTCAAGTGGC TCAATGGAGT AGTAGGC 27

<210> 28
<211> 85
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 28
GCCAGCCCCC GATTGGGGGC GACACTCCAC CATAGATCAC TCCCCTGTGA GGAACTACTG 60
TCTTCACGCA GAAAGCGTCT AGCCA 85

<210> 29
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 29
GACCACGCG TATCGATGT CGACTTTTTT TTTTTTTTTT 40

<210> 30

<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 30
GACCACGCGT ATCGATGTCG AC 22

<210> 31
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 31
ATGGACCAGG AGTTGGCTGC ATCGTGC 27

<210> 32
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial sequence: PCR primer

<400> 32
CTAACGCGCA CGCACGAATG AGGCCTT 27

<210> 33
<211> 9609
<212> DNA
<213> Hepatitis C virus

<400> 33
gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgtag cctccgggac 120
ccccccctcc gggagagcca tggtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgccccc 240
gcgagactac tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg attcctaaac 360
ctcaaagaaa aaccaaacgt aacacctacc gccgcccaca ggacgtcaag ttcccgggcg 420
gtggccagat cgttggtgga gtttacctgt tgccgcgcag gggccccagg ttgggtgtgc 480
gcgcgactag gaagacttcc gagcggtcgc aacctcgtgg aaggcgacag cctatcccca 540
aggctcgccg gcccgagggc tgggcctggg cacaacccgg gtacccttgg cccctctatg 600
gcaatgaggg cttggggtgg gcaggatggc tcctgtcacc ccgtggctcc cgccctagtt 660
ggggccccat agacccccgg cgtaggtcgc gtaatttggg taaggtcatc gataccctca 720

catgcggctt cgccgacctc atgggataca ttccgctcgt cggcgccccc ttaggggcg 780
ctgccagggc cctggcgcat ggcgtcaggg ttctggagga cggcgtgaac tatgcaacag 840
ggaaccttcc cggttgctct ttctctatct tcctcttggc tttgctgtct tgtttgacca 900
tcccagcttc cgcttatgaa gtgcgcaacg cgtccggggt gtaccatgtt acgaacgact 960
gctccaactc aagcattgtg tatgaggcag cggacataat catgcacacc cccggatgcg 1020
taccctgcgt tcgggagaac aactcttccc gctgctgggt agcgctcact cccacgctcg 1080
cggccaggaa cctcagcgtc cccaccacga caatacgacg ccatgtcgat ttgctcgttg 1140
gggcggctgc tttctgctcc gctatgtacg tgggggatct ttgcggatcc gttctactcg 1200
tctcccagct gttcaccctc tcgcctcgcc agcatgagac agtacaagat tgcaattgct 1260
cactttatcc cggccacgta acaggtcacc gcatggcttg ggatatgatg atgaattggt 1320
cgcctacaac agccctagtg ttatcgcaga tactccggat cccacaaact atcgtggaca 1380
tggtggcggg ggcccactgg ggagtcctgg cgggcattgc ctactattcc atggtgggga 1440
actgggctaa ggttttggtt gtgatgctac tctttgctgg cgttgacggg cacacccaag 1500
tgatgggggg gtcgcaagcc tctaccatca acacgctcac gggaattttc tcccccgggg 1560
cgaaacagaa gattcagctc ataaacacca acggcagttg gcacatcaac aggactgccc 1620
tgaactgcaa tgactccctc aacactgggt tccttgccgc gctgttctac acccacagct 1680
tcaactcgtc cggatgccta gagcgtatgg ccagctgccg ccccattgac aagttcgatc 1740
aggggtgggg tcccattact tatgatgaag gccccgattt agaccagagg ccctattgct 1800
ggcactacgc accccggtcg tgcggtattg tgcccgcgtc gcaggtgtgt ggcccagtgt 1860
attgtttcac cccaagccct gtcgtggtgg ggacgaccga tcgtaccggt gcccccacgt 1920
atagatgggg ggagaatgag acggacgtgc tgatcctcaa caacacgcgg ccgccgcagg 1980
gcaactggtt cggctgcaca tggatgaata gcactgggtt caccaagaca tgcgggggcc 2040
ccccatgtaa catcgggggg gctggtaaca acaccttggt ctgccccacg gactgcttcc 2100
ggaaacaccc cgaggctact tacaccaaat gcggttcggg accctggttg acacctaggt 2160
gcatggttga ctacccatac aggccttggc actacccctg tactgtcaac tttaccattt 2220
tcaaggttag gatgtatgtg gggggcgtag agcacaggct caacgctgca cgcaactgga 2280
ctcgaggaga gcgttgtgac ttggaggaca gggatagatc agagcttagc ccgctgctgc 2340
tgtctacgac agagtggcag atactgcctt gctccttcac cactctaccg gctctgtcca 2400
ctggcttgat ccacctccat cagaacatcg tggacgtgca gtatctgtac ggtatagggt 2460
cagtggttgt ctccctcgtg atcaaatggg agtatgtctt gctgtccttc ttcctcctgg 2520
cggacgcgcg cgtctgtgcc tgtttgtgga tgatgctgct gatagcccag gctgaggccg 2580
ccttagagaa cctggtggtc ctcaatgcgg cgtccgtggc cggagcgcat ggcatcctca 2640
ccttccttgt gttcttctgt gctgcctggt acatcaaggg caggctggtc cctggggcgg 2700
catatgcttt ttacggcgtg tggccgctgc tcctgctctt gctggcgctg ccaccacgag 2760
cgtacgctat ggaccgggag atggctgcat cgtgcggagg cgcggttttc ataggtctgg 2820
tactcctgac cctgtcacca cactataaag tgttccttgc taggctcata tggtggttgc 2880
agtactttac caccagagcc gaggcaattt tgcacgtgtg ggtccccct ctcaacgtcc 2940
gagggggccg tgatgccgtc attctcctca cgtgcgcggt ccacccagac ctaatctttg 3000
acatcaccaa actcttgctc gctgtgctcg gtccgctcat ggtgtttctg gccggcataa 3060
caagggtgcc atacttcgtg cgcgcccaag gcctcattcg tgcgtgcgcg ttagcgcgga 3120
aggtcgctgg gggtcattat atccaaatgg ctctcatgaa gctggccgcg ctgacaggca 3180
cgtatcttta cgaccatctt actccactgc gggactgggc ccatgcgggc ctgcgagacc 3240
ttgcggtagc agtcgagcct gtcgtctttt ctgacatgga gaccaagatc atcacctggg 3300
gggcagacac cgcagcgtgt ggggacatca tcctgggtct gcctgtctcc gctcgaaggg 3360
ggagggagat acttttggga ccagccgata gtcttgaagg gcaggggtgg cgactccttg 3420
cgcctattac ggcctactcc cagcagacgc ggggcctact tggctgcatc atcactagtc 3480
tcacaggtcg ggacaagaac caggtcgagg gggaggttca agtggtttcc accgcaacac 3540
agtctttcct agcgacctgc gtcaacggcg tgtgttggac tgtcttccat ggtgccggct 3600
caaaaacctt agccggtcca aaggggccaa ttactcaaat gtacaccaac gtggaccagg 3660

acctcgtcgg ctggcaggcg ccccccgggg cgcgttcctt gacaccatgc acctgcggca 3720
gctcggacct ttacctggtc acgaggcacg ctgacgtcat tccggtgcgc cggcggggcg 3780
acaacagagg gagcctactc tcccccaggc ccgtctccta cttaaagggc tcttcgggtg 3840
gtccactcct ctgcccctcg gggcatgccg taggcatctt ccgggccgcc gtatgcaccc 3900
gggggggttgc gaaggcggtg gactttatac ccgttgagtc tatggagact accatgcggt 3960
ccccggtctt cacagacaat tcatcccccc cggccgtacc gcagacattc caagtggctc 4020
atctacacgc tcccactggc agcggtaaga gcactaaggt gccggctgcg tatgcagccc 4080
aagggtacaa ggtacttgtc ctgaacccgt ccgtggccgc taccttaggg ttcggaacgt 4140
atatgtctaa ggcatatggt gtcgatccta acatcagaac cggggtaagg accatcacca 4200
cgggcgcccc catcacgtac tccacttacg gcaagtttct tgccgacggt ggttgctccg 4260
ggggagccta cgacatcata atgtgtgatg aatgccactc aactgactca actactgtct 4320
tgggcatcgg cacagtcctg gaccaagcgg agacggctgg agcgcggctt gtcgtgctcg 4380
ccaccgctac gcctccggga tcggtcaccg tgccacatcc caatattgag gagatcgctc 4440
tatccaacac tggagagatc cccttctatg gcaaagccat ccccattgag accatcaagg 4500
gggggaggca tctcattttc tgccattcca agaagaaatg tgacgagctc gctgcaaagt 4560
tgtcgggcct cggacttaac gctgtagcgt attaccgggg ccttgatgtg tccgttatac 4620
caactagcgg agacgtcgtt gtcgtggcaa cagacgctct aatgacgggc tttaccggcg 4680
attccgactc agtgatcgac tgtaatacat gtgtcaccca gacagttgat ttcagcttgg 4740
accctacctt caccattgag acgacgaccg tgccccaaga cgcggtgtcg cgctcgcagc 4800
ggcgaggcag gactggtagg ggtaggggggg gcatctacag gtttgtgact ccaggagaac 4860
ggcccctcggg catgttcgat tcttcggtcc tgtgtgagtg ctatgacgcg ggctgtgctt 4920
ggtacgagct cacgcccgcc gagacctcag ttaggttgcg ggcttacctg aatacaccag 4980
ggttacccgt ctgccaggac catctggagt tctgggagag tgtcttcaca ggcctcaccc 5040
acatagacgc ccacttcttg tctcagacta agcaggcagg agacaacttc ccctacctgg 5100
tagcgtacca ggctacagtg tgcgccagag cgcaggctcc acctccatca tgggatcaaa 5160
tgtggaagtc tctcatacgg ctaaagccta cactgcacgg gccaacaccc ctgctatata 5220
ggctagggac cgtccaaagc gagattaccc tcacacaccc agtgacaaaa tacatcatgg 5280
catgcatgtc ggctgacctg gaggtcgtca cgagcacctg ggtgctagtg agcggagtcc 5340
tcgcagccct ggccgcgtac tgcttgacaa caggcagcgt ggtcattgtg ggcaggatca 5400
tcttgtccgg gaagcctgct gttattcccg acagggaagt tctttaccgg gagttcgatg 5460
agatggagga gtgcgcctca cacctccctt acatcgaaca ggggatgcag ctcgctgagc 5520
agttcaaaca aaaggcgctt gggctgctgc aaacggccac caagcaagcg gaggctgctg 5580
ctcccgtggt ggagtccaag tggcgggccc ttgaggcctt ttgggcgaag cacatgtgga 5640
acttcatcag cggaatacag tacctggcag gcttgtccac tctgcatggc aatcccgcaa 5700
tagcatcact gatggcattt acagcctcta tcaccagccc gctcaccacc caacacaccc 5760
tcctgtttaa catcttgggg ggatgggtgg ccgcccaact tgctcccccc agtgctgctt 5820
cggctttcgt aggtgccggc atcgctggtg cagccgtcgg cagcataggt ctcgggaagg 5880
tgctcgttga cgtcctggcg ggttatggag caggcgtggc aggcgcactt gtggccttta 5940
aggttatgag tggcgaggtg ccctccaccg aggacctggt caatttgctc cctgccgtcc 6000
tctctcctgg tgccctggtc gtcggggtcg tgtgcgcagc aatactgcgt cggcacgtgg 6060
gcccaggaga gggggctgtg cagtggatga accggctgat agcgttcgct tcgcggggta 6120
accatgtttc ccccacgcac tatgtgcctg aaagcgacgc cgcagcacgt gtcactcaaa 6180
tcctctccag cctcaccgtc acccaactgt tgaggaggct ccaccaatgg attaatgagg 6240
attgctccac gccgtgctcc ggttcctggc tgagagatgt ttgggactgg atatgtacgg 6300
tgctgactga cttcaagacc tggctccaat ccaaactcct gccgcggtta ccaggagtcc 6360
ccttttttctc gtgtcaacgc gggtacaggg gagtctggcg gggagacggc atcatgcaga 6420
ccacctgccc atgtggagca caaatctctg gacatgtcaa aaatggttcc atgaggatcg 6480
ttgggcctag atcctgcagc aacacgtggc atggaacatt ccccattaat gcgtacacca 6540
cgggcccctg cacgcccgcc ccggcgccaa actactccag ggcactgtgg cgggtggctg 6600

ctgaggagta cgtggaggtc acgcgggtgg gggatttcca ctacgtgaca ggcatgacca 6660
ctgacaacgt aaagtgccca tgtcaggttc cagcccccga attcttcaca gaggtggatg 6720
gagtgcggct gcacaggtac gctccggcgt gcaaacccct cctacgggag gaggtcacat 6780
tccaggtcgg gctcaaccag tacttggttg ggtcacagct cccgtgtgag cccgagccgg 6840
acgtagcagt gctcacctcc atgcttaccg atccttctca catcacagca gagacggcta 6900
agcgtaggct ggataggggg tctcccccct ccttggccag ttcttcagcc agccagttat 6960
ctgcgccttc tttgaaggcg acttgcacta cccgtcacga ctccccagac gccggcctca 7020
ttgaggccaa tctcctgtgg cggcaggaga tgggcgggaa catcactcgc gtggagtcag 7080
agaacaaggt agtgattctg gactctttcg agccgcttcg agcggaagag gatgagaggg 7140
aagtatccgt tccggcggaa atcctgcgaa aatccaggaa attccccaga gcgatgccca 7200
tatgggcacg cccggattac aatcctccac tgctagagtc ctggaagaac ccggactacg 7260
tccctccggt ggtacacggg tgcccattgc cgcccatcaa gggcccccca ataccacccc 7320
cacggagaaa gaggacggtt gtcctaacag aatccaccgt gtcttctgcc ttggcggagc 7380
tcgccacaaa gaccttcggc agctccggat catcggctgt cgacagcggt acagcgagcg 7440
ccccccctga ccagccctct gacaacggcg acgcgggatc cgacgctgag tcgtactcct 7500
ccatgccccc ccttgaaggg gagccggggg accccgatct cagcgacggg tcttggtcta 7560
ccgtgagcga ggaggccagt gaggacgtcg tctgctgctc gatgtcctac acgtggacag 7620
gcgctctgat cacaccatgc gctgcggagg agagcaaatt acccatcaat gcgctgagca 7680
actccttgct gcgtcaccac aacatggtct atgccacaac atcccgcagt gccagtcagc 7740
gacagaagaa ggttaccttt gacagactgc aagtcctgga cgaccactac cgggacgtgc 7800
ttaaggagat gaaggcgaag gcgtccacag tcaaggccaa gcttctatcc gtagaagaag 7860
cctgtaagct gacgccccca cactcggcca aatccaaatt tggctacggg gcaaaggacg 7920
tccggaacct gtccagcaag gccgttaacc acatccgctc cgtgtggaag gacttgctgg 7980
aagacactga gacaccaata gacaccacca tcatggcaaa gaacgaggtc ttctgcgtcc 8040
agccagagaa gggaggccgc aagccagctc gccttatcgt gttcccggat ttgggggtcc 8100
gtgtgtgcga gaaaatggcc ctttacgacg tggtctccac cctccctcag gccgtgatgg 8160
gctcctcata cggattccag tactctcctg gacagcgggt cgagttcctg gtgaatgcct 8220
ggaaatcaaa gaaaagcccc atgggcttcg catatgacac ccgctgtttt gactcaacgg 8280
tcactgagag tgacatccgt gttgaggagt caatttacca atgttgtgac ttagtccccg 8340
aagctagaca ggccataagg tcgctcacag agcgacttta tgtcgggggt cctctgacta 8400
attcaaaagg gcagaactgc ggttatcgcc ggtgccgcgc gagcggcgtg ctgacgacta 8460
gctgcggtaa cacccttaca tgttacttga aggcctctgc agcctgtcga gctgcaaagc 8520
tccaggactg cacgatgctc gtgtgcggag acgaccttgt cgttatctgt gaaagcgcgg 8580
gaacccaaga ggacgcggcg agcctacgag tctttacgga ggctatgact aggtattccg 8640
ccccccccgg ggacccgccc caaccagaat acgacttgga gctgataaca tcgtgctcct 8700
ccaatgtgtc ggtcgctcac gatgcatctg gcaaaagagt gtactacctc acccgtgacc 8760
ccaccacccc ccttgcgcgg gctgcgtggg agacagctag acacactcca gtcaactcct 8820
ggctaggcaa catcattatg tacgcaccca ctttgtgggc aaggatgatt ctgatgactc 8880
acttcttctc catcctccta gcccaggagc aacttgaaaa agccctagat tgtcaaatct 8940
acggggccta ctactccatt gagccacttg acctacctca gatcattgaa cgactccacg 9000
gtcttagcgc attttcgctc catagttact ctccaggcga gatcaatagg gtggcttcat 9060
gcctcagaaa acttggggta ccacccttgc gagcctggag acatcgggcc aggagtgtcc 9120
gcgctaagct gctgtctcag gggggaaggg ccgccacttg tggcaaatac ctcttcaact 9180
gggcagtaaa gaccaagctc aaactcactc caattccggc tgcgtcccag ttggacttgt 9240
ccggctggtt cgttgctggc tacggcgggg gagacatata tcacagcctg tctcgcgccc 9300
gaccccgctg gttcatgttg tgcctacccc tactctctgt aggggtaggc atcaacctgc 9360
tccccaaccg gtgaatgggg agctaaacac tccaggccaa taggccaacc tgttttttttt 9420
ttttttttt ttttttttcc ttttttttt ttttttttt ttttttttt ttttccttt 9480
tttctttttt ttttttttt tccttccttt tggtggctcc atcttagccc tagtcacggc 9540

tagctgtgaa aggtccgtga gccgcatgac tgcagagagt gctgatactg gcctctctgc 9600
agatcatgt 9609

<210> 34
<211> 3010
<212> PRT
<213> Hepatitis C virus

<400> 34

Met Ser Thr Ile Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Tyr
1 5 10 15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
35 40 45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
50 55 60

Ile Pro Lys Ala Arg Arg Pro Glu Gly Trp Ala Trp Ala Gln Pro Gly
65 70 75 80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
85 90 95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Ile Asp Pro
100 105 110

Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys
115 120 125

Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu
130 135 140

Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
145 150 155 160

Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
165 170 175

Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr
180 185 190

Glu Val Arg Asn Ala Ser Gly Val Tyr His Val Thr Asn Asp Cys Ser
195 200 205

```
Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Ile Ile Met His Thr Pro
    210                 215                 220

Gly Cys Val Pro Cys Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val
225                 230                 235                 240

Ala Leu Thr Pro Thr Leu Ala Ala Arg Asn Leu Ser Val Pro Thr Thr
                245                 250                 255

Thr Ile Arg Arg His Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys
            260                 265                 270

Ser Ala Met Tyr Val Gly Asp Leu Cys Gly Ser Val Leu Leu Val Ser
        275                 280                 285

Gln Leu Phe Thr Leu Ser Pro Arg Gln His Glu Thr Val Gln Asp Cys
    290                 295                 300

Asn Cys Ser Leu Tyr Pro Gly His Val Thr Gly His Arg Met Ala Trp
305                 310                 315                 320

Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val Leu Ser Gln
                325                 330                 335

Ile Leu Arg Ile Pro Gln Thr Ile Val Asp Met Val Ala Gly Ala His
            340                 345                 350

Trp Gly Val Leu Ala Gly Ile Ala Tyr Tyr Ser Met Val Gly Asn Trp
        355                 360                 365

Ala Lys Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His
    370                 375                 380

Thr Gln Val Met Gly Gly Ser Gln Ala Ser Thr Ile Asn Thr Leu Thr
385                 390                 395                 400

Gly Ile Phe Ser Pro Gly Ala Lys Gln Lys Ile Gln Leu Ile Asn Thr
                405                 410                 415

Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
            420                 425                 430

Leu Asn Thr Gly Phe Leu Ala Ala Leu Phe Tyr Thr His Ser Phe Asn
        435                 440                 445

Ser Ser Gly Cys Leu Glu Arg Met Ala Ser Cys Arg Pro Ile Asp Lys
    450                 455                 460
```

```
Phe Asp Gln Gly Trp Gly Pro Ile Thr Tyr Asp Glu Gly Pro Asp Leu
465                 470                 475                 480

Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Ser Cys Gly Ile
                485                 490                 495

Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser
            500                 505                 510

Pro Val Val Val Gly Thr Thr Asp Arg Thr Gly Ala Pro Thr Tyr Arg
        515                 520                 525

Trp Gly Glu Asn Glu Thr Asp Val Leu Ile Leu Asn Asn Thr Arg Pro
    530                 535                 540

Pro Gln Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe
545                 550                 555                 560

Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn
                565                 570                 575

Asn Thr Leu Val Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala
            580                 585                 590

Thr Tyr Thr Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met
        595                 600                 605

Val Asp Tyr Pro Tyr Arg Pro Trp His Tyr Pro Cys Thr Val Asn Phe
    610                 615                 620

Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Leu
625                 630                 635                 640

Asn Ala Ala Arg Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp
                645                 650                 655

Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp
            660                 665                 670

Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly
        675                 680                 685

Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly
    690                 695                 700

Ile Gly Ser Val Val Val Ser Leu Val Ile Lys Trp Glu Tyr Val Leu
705                 710                 715                 720
```

```
Leu Ser Phe Phe Leu Leu Ala Asp Ala Arg Val Cys Ala Cys Leu Trp
                725                 730                 735

Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val
            740                 745                 750

Val Leu Asn Ala Ala Ser Val Ala Gly Ala His Gly Ile Leu Thr Phe
        755                 760                 765

Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro
    770                 775                 780

Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu
785                 790                 795                 800

Leu Ala Leu Pro Pro Arg Ala Tyr Ala Met Asp Arg Glu Met Ala Ala
                805                 810                 815

Ser Cys Gly Gly Ala Val Phe Ile Gly Leu Val Leu Leu Thr Leu Ser
            820                 825                 830

Pro His Tyr Lys Val Phe Leu Ala Arg Leu Ile Trp Trp Leu Gln Tyr
        835                 840                 845

Phe Thr Thr Arg Ala Glu Ala Ile Leu His Val Trp Val Pro Pro Leu
    850                 855                 860

Asn Val Arg Gly Gly Arg Asp Ala Val Ile Leu Leu Thr Cys Ala Val
865                 870                 875                 880

His Pro Asp Leu Ile Phe Asp Ile Thr Lys Leu Leu Leu Ala Val Leu
                885                 890                 895

Gly Pro Leu Met Val Phe Leu Ala Gly Ile Thr Arg Val Pro Tyr Phe
            900                 905                 910

Val Arg Ala Gln Gly Leu Ile Arg Ala Cys Ala Leu Ala Arg Lys Val
        915                 920                 925

Ala Gly Gly His Tyr Ile Gln Met Ala Leu Met Lys Leu Ala Ala Leu
    930                 935                 940

Thr Gly Thr Tyr Leu Tyr Asp His Leu Thr Pro Leu Arg Asp Trp Ala
945                 950                 955                 960

His Ala Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe
                965                 970                 975
```

Ser Asp Met Glu Thr Lys Ile Ile Thr Trp Gly Ala Asp Thr Ala Ala
980 985 990

Cys Gly Asp Ile Ile Leu Gly Leu Pro Val Ser Ala Arg Arg Gly Arg
995 1000 1005

Glu Ile Leu Leu Gly Pro Ala Asp Ser Leu Glu Gly Gln Gly Trp Arg
1010 1015 1020

Leu Leu Ala Pro Ile Thr Ala Tyr Ser Gln Gln Thr Arg Gly Leu Leu
1025 1030 1035 1040

Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu
1045 1050 1055

Gly Glu Val Gln Val Val Ser Thr Ala Thr Gln Ser Phe Leu Ala Thr
1060 1065 1070

Cys Val Asn Gly Val Cys Trp Thr Val Phe His Gly Ala Gly Ser Lys
1075 1080 1085

Thr Leu Ala Gly Pro Lys Gly Pro Ile Thr Gln Met Tyr Thr Asn Val
1090 1095 1100

Asp Gln Asp Leu Val Gly Trp Gln Ala Pro Pro Gly Ala Arg Ser Leu
1105 1110 1115 1120

Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His
1125 1130 1135

Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Asn Arg Gly Ser Leu
1140 1145 1150

Leu Ser Pro Arg Pro Val Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro
1155 1160 1165

Leu Leu Cys Pro Ser Gly His Ala Val Gly Ile Phe Arg Ala Ala Val
1170 1175 1180

Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Ser
1185 1190 1195 1200

Met Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro
1205 1210 1215

Pro Ala Val Pro Gln Thr Phe Gln Val Ala His Leu His Ala Pro Thr
1220 1225 1230

Gly Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly
1235 1240 1245

Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe
1250 1255 1260

Gly Thr Tyr Met Ser Lys Ala Tyr Gly Val Asp Pro Asn Ile Arg Thr
1265 1270 1275 1280

Gly Val Arg Thr Ile Thr Thr Gly Ala Pro Ile Thr Tyr Ser Thr Tyr
1285 1290 1295

Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile
1300 1305 1310

Ile Met Cys Asp Glu Cys His Ser Thr Asp Ser Thr Thr Val Leu Gly
1315 1320 1325

Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val
1330 1335 1340

Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro
1345 1350 1355 1360

Asn Ile Glu Glu Ile Ala Leu Ser Asn Thr Gly Glu Ile Pro Phe Tyr
1365 1370 1375

Gly Lys Ala Ile Pro Ile Glu Thr Ile Lys Gly Gly Arg His Leu Ile
1380 1385 1390

Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Ser
1395 1400 1405

Gly Leu Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser
1410 1415 1420

Val Ile Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu
1425 1430 1435 1440

Met Thr Gly Phe Thr Gly Asp Ser Asp Ser Val Ile Asp Cys Asn Thr
1445 1450 1455

Cys Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile
1460 1465 1470

Glu Thr Thr Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg
1475 1480 1485

Gly Arg Thr Gly Arg Gly Arg Gly Gly Ile Tyr Arg Phe Val Thr Pro
1490 1495 1500

Gly Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys
1505 1510 1515 1520

Tyr Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Ser
1525 1530 1535

Val Arg Leu Arg Ala Tyr Leu Asn Thr Pro Gly Leu Pro Val Cys Gln
1540 1545 1550

Asp His Leu Glu Phe Trp Glu Ser Val Phe Thr Gly Leu Thr His Ile
1555 1560 1565

Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Asp Asn Phe Pro
1570 1575 1580

Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro
1585 1590 1595 1600

Pro Pro Ser Trp Asp Gln Met Trp Lys Ser Leu Ile Arg Leu Lys Pro
1605 1610 1615

Thr Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Thr Val Gln
1620 1625 1630

Ser Glu Ile Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Ala Cys
1635 1640 1645

Met Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Ser
1650 1655 1660

Gly Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Thr Thr Gly Ser Val
1665 1670 1675 1680

Val Ile Val Gly Arg Ile Ile Leu Ser Gly Lys Pro Ala Val Ile Pro
1685 1690 1695

Asp Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Cys Ala
1700 1705 1710

Ser His Leu Pro Tyr Ile Glu Gln Gly Met Gln Leu Ala Glu Gln Phe
1715 1720 1725

Lys Gln Lys Ala Leu Gly Leu Leu Gln Thr Ala Thr Lys Gln Ala Glu
1730 1735 1740

```
Ala Ala Ala Pro Val Val Glu Ser Lys Trp Arg Ala Leu Glu Ala Phe
1745                1750                1755                1760

Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala
                1765                1770                1775

Gly Leu Ser Thr Leu His Gly Asn Pro Ala Ile Ala Ser Leu Met Ala
            1780                1785                1790

Phe Thr Ala Ser Ile Thr Ser Pro Leu Thr Thr Gln His Thr Leu Leu
        1795                1800                1805

Phe Asn Ile Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Pro Pro Ser
    1810                1815                1820

Ala Ala Ser Ala Phe Val Gly Ala Gly Ile Ala Gly Ala Ala Val Gly
1825                1830                1835                1840

Ser Ile Gly Leu Gly Lys Val Leu Val Asp Val Leu Ala Gly Tyr Gly
                1845                1850                1855

Ala Gly Val Ala Gly Ala Leu Val Ala Phe Lys Val Met Ser Gly Glu
            1860                1865                1870

Val Pro Ser Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Val Leu Ser
        1875                1880                1885

Pro Gly Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg
    1890                1895                1900

His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile
1905                1910                1915                1920

Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro
                1925                1930                1935

Glu Ser Asp Ala Ala Ala Arg Val Thr Gln Ile Leu Ser Ser Leu Thr
            1940                1945                1950

Val Thr Gln Leu Leu Arg Arg Leu His Gln Trp Ile Asn Glu Asp Cys
        1955                1960                1965

Ser Thr Pro Cys Ser Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Ile
    1970                1975                1980

Cys Thr Val Leu Thr Asp Phe Lys Thr Trp Leu Gln Ser Lys Leu Leu
1985                1990                1995                2000
```

Pro Arg Leu Pro Gly Val Pro Phe Phe Ser Cys Gln Arg Gly Tyr Arg
2005 2010 2015

Gly Val Trp Arg Gly Asp Gly Ile Met Gln Thr Thr Cys Pro Cys Gly
2020 2025 2030

Ala Gln Ile Ser Gly His Val Lys Asn Gly Ser Met Arg Ile Val Gly
2035 2040 2045

Pro Arg Ser Cys Ser Asn Thr Trp His Gly Thr Phe Pro Ile Asn Ala
2050 2055 2060

Tyr Thr Thr Gly Pro Cys Thr Pro Ala Pro Ala Pro Asn Tyr Ser Arg
2065 2070 2075 2080

Ala Leu Trp Arg Val Ala Ala Glu Glu Tyr Val Glu Val Thr Arg Val
2085 2090 2095

Gly Asp Phe His Tyr Val Thr Gly Met Thr Thr Asp Asn Val Lys Cys
2100 2105 2110

Pro Cys Gln Val Pro Ala Pro Glu Phe Phe Thr Glu Val Asp Gly Val
2115 2120 2125

Arg Leu His Arg Tyr Ala Pro Ala Cys Lys Pro Leu Leu Arg Glu Glu
2130 2135 2140

Val Thr Phe Gln Val Gly Leu Asn Gln Tyr Leu Val Gly Ser Gln Leu
2145 2150 2155 2160

Pro Cys Glu Pro Glu Pro Asp Val Ala Val Leu Thr Ser Met Leu Thr
2165 2170 2175

Asp Pro Ser His Ile Thr Ala Glu Thr Ala Lys Arg Arg Leu Asp Arg
2180 2185 2190

Gly Ser Pro Pro Ser Leu Ala Ser Ser Ser Ala Ser Gln Leu Ser Ala
2195 2200 2205

Pro Ser Leu Lys Ala Thr Cys Thr Thr Arg His Asp Ser Pro Asp Ala
2210 2215 2220

Gly Leu Ile Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn
2225 2230 2235 2240

Ile Thr Arg Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe
2245 2250 2255

Glu Pro Leu Arg Ala Glu Glu Asp Glu Arg Glu Val Ser Val Pro Ala
2260 2265 2270

Glu Ile Leu Arg Lys Ser Arg Lys Phe Pro Arg Ala Met Pro Ile Trp
2275 2280 2285

Ala Arg Pro Asp Tyr Asn Pro Pro Leu Leu Glu Ser Trp Lys Asn Pro
2290 2295 2300

Asp Tyr Val Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Ile Lys
2305 2310 2315 2320

Gly Pro Pro Ile Pro Pro Pro Arg Arg Lys Arg Thr Val Val Leu Thr
2325 2330 2335

Glu Ser Thr Val Ser Ser Ala Leu Ala Glu Leu Ala Thr Lys Thr Phe
2340 2345 2350

Gly Ser Ser Gly Ser Ser Ala Val Asp Ser Gly Thr Ala Ser Ala Pro
2355 2360 2365

Pro Asp Gln Pro Ser Asp Asn Gly Asp Ala Gly Ser Asp Ala Glu Ser
2370 2375 2380

Tyr Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu
2385 2390 2395 2400

Ser Asp Gly Ser Trp Ser Thr Val Ser Glu Glu Ala Ser Glu Asp Val
2405 2410 2415

Val Cys Cys Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Ile Thr Pro
2420 2425 2430

Cys Ala Ala Glu Glu Ser Lys Leu Pro Ile Asn Ala Leu Ser Asn Ser
2435 2440 2445

Leu Leu Arg His His Asn Met Val Tyr Ala Thr Thr Ser Arg Ser Ala
2450 2455 2460

Ser Gln Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu Asp
2465 2470 2475 2480

Asp His Tyr Arg Asp Val Leu Lys Glu Met Lys Ala Lys Ala Ser Thr
2485 2490 2495

Val Lys Ala Lys Leu Leu Ser Val Glu Glu Ala Cys Lys Leu Thr Pro
2500 2505 2510

Pro His Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val Arg
2515 2520 2525

Asn Leu Ser Ser Lys Ala Val Asn His Ile Arg Ser Val Trp Lys Asp
2530 2535 2540

Leu Leu Glu Asp Thr Glu Thr Pro Ile Asp Thr Thr Ile Met Ala Lys
2545 2550 2555 2560

Asn Glu Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro Ala
2565 2570 2575

Arg Leu Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys Met
2580 2585 2590

Ala Leu Tyr Asp Val Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser
2595 2600 2605

Ser Tyr Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val
2610 2615 2620

Asn Ala Trp Lys Ser Lys Lys Ser Pro Met Gly Phe Ala Tyr Asp Thr
2625 2630 2635 2640

Arg Cys Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Val Glu Glu
2645 2650 2655

Ser Ile Tyr Gln Cys Cys Asp Leu Val Pro Glu Ala Arg Gln Ala Ile
2660 2665 2670

Arg Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn Ser
2675 2680 2685

Lys Gly Gln Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu
2690 2695 2700

Thr Thr Ser Cys Gly Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala
2705 2710 2715 2720

Ala Cys Arg Ala Ala Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly
2725 2730 2735

Asp Asp Leu Val Val Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala
2740 2745 2750

Ala Ser Leu Arg Val Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro
2755 2760 2765

```
Pro Gly Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser
   2770                2775                2780

Cys Ser Ser Asn Val Ser Val Ala His Asp Ala Ser Gly Lys Arg Val
2785                2790                2795                2800

Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala Trp
                2805                2810                2815

Glu Thr Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile
           2820                2825                2830

Met Tyr Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His Phe
       2835                2840                2845

Phe Ser Ile Leu Leu Ala Gln Glu Gln Leu Glu Lys Ala Leu Asp Cys
   2850                2855                2860

Gln Ile Tyr Gly Ala Tyr Tyr Ser Ile Glu Pro Leu Asp Leu Pro Gln
2865                2870                2875                2880

Ile Ile Glu Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser Tyr
                2885                2890                2895

Ser Pro Gly Glu Ile Asn Arg Val Ala Ser Cys Leu Arg Lys Leu Gly
           2900                2905                2910

Val Pro Pro Leu Arg Ala Trp Arg His Arg Ala Arg Ser Val Arg Ala
       2915                2920                2925

Lys Leu Leu Ser Gln Gly Gly Arg Ala Ala Thr Cys Gly Lys Tyr Leu
   2930                2935                2940

Phe Asn Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Ile Pro Ala
2945                2950                2955                2960

Ala Ser Gln Leu Asp Leu Ser Gly Trp Phe Val Ala Gly Tyr Gly Gly
                2965                2970                2975

Gly Asp Ile Tyr His Ser Leu Ser Arg Ala Arg Pro Arg Trp Phe Met
           2980                2985                2990

Leu Cys Leu Pro Leu Leu Ser Val Gly Val Gly Ile Asn Leu Leu Pro
       2995                3000                3005

Asn Arg
   3010
```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 January 2002 (31.01.2002)

PCT

(10) International Publication Number
WO 02/008447 A3

(51) International Patent Classification[7]: C12Q 1/70

(21) International Application Number: PCT/IL01/00669

(22) International Filing Date: 20 July 2001 (20.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/220,248 24 July 2000 (24.07.2000) US

(71) Applicant *(for all designated States except US)*: **INSTITUTE OF MOLECULAR & CELL BIOLOGY** [SG/SG]; 30 Medical Drive, 117609 Singapore (SG).

(71) Applicant *(for BB only)*: **EHRLICH, Gal** [IL/IL]; 28 Bezalel Street, 52 521 Ramat Gan (IL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TAN, Yin, Hwee** [SG/SG]; 3 Pandan Valley # 11-308, Chempaka Court, 597627 Singapore (SG). **LIM, Siew, Pheng** [SG/SG]; 427 #04-314 Serangoon Central, 550427 Singapore (SG). **LIM, Seng, Gee** [SG/SG]; 113 Clementi Road # 04-08, 129789 Singapore (SG). **HONG, Wan, Jin** [SG/SG]; 835 Bukit Timah Road, #01-07 Royal Ville, 279888 Singapore (SG).

(74) Agent: **G. E. EHRLICH (1995) LTD.**; 28 Bezalel Street, 52 521 Ramat Gan (IL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
30 October 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NUCLEIC ACIDS AND METHODS FOR DETECTING VIRAL INFECTION, UNCOVERING ANTI-VIRAL DRUG CANDIDATES AND DETERMINING DRUG RESISTANCE OF VIRAL ISOLATES

(57) Abstract: A nucleic acid construct is provided. The nucleic acid construct includes (a) an expression cassette including: (i) a first polynucleotide region including a 5' NCR sequence of an RNA virus and at least an N-terminal portion of a core sequence of the RNA virus; (ii) a second polynucleotide region including a 3' UTR sequence of the RNA virus and at least a C-terminal portion of a polymerase sequence of the virus; and (iii) a third polynucleotide region encoding a reporter molecule, the third polynucleotide region being flanked by the first and the second polynucleotide regions; and (b) a promoter sequence being operatively linked to the expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette.

WO 02/008447 A3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IL01/00669

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/70
US CL : 536/ 23.1; 435/ 5, 6, 69.1, 173.3, 235.1, 320.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 536/ 23.1; 435/ 5, 6, 69.1, 173.3, 235.1, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Lohmann, V. Replication of Subgenomic Hepatitis C Virus in a Hepatoma Cell Line. Science 1999 Vol. 285, pages 110- 113, see Figure 1.	1, 4- 7- 15, 18- 20, 23- 25, 28, 33, and 36- 41
Y	VARNAVSKI, A.N., Stable High-Level Expression of Heterologous Genes IN Vitro and In Vivo by Noncytopathic DNA-Based Kunjin Virus Replication System Journal Of Virology May 2000, Vol 74, No. 9, pages 4394- 4403, see Figures 1, and 2 and Discussion.	1, 4- 7- 15, 18- 20, 23- 25, 28, 33, and 36- 41
Y	KHROMYKH, A.A. et al. cis- and trans- Acting Elements in Flavivirus RNA Replication Journal of Virology April 200 Vol 74, No. 7, pages 3253- 3263, see Figures 1 and 6, Table 1, and Discussion.	1, 4- 7- 15, 18- 20, 23- 25, 28, 33, and 36- 41
A	Ikeda, M. et al. Selectable Subgenomic and Genome-Length Dicistronic RNAs Derived from an Infectious Molecular Clone of the HCV-N Strain of Hepatitis C Virus Replicates Efficiently in Cultured Huh7 Cells.Journal of Virology 2002 Vol. 76, pages 2997-3006, Figure 1A, 6	1, 4- 7- 15, 18- 20, 23- 25, 28, 33, and 36- 41

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
21 March 2003 (21.03.2003)

Date of mailing of the international search report
29 AUG 2003

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Myron G. Hill
James C. Housel
for
Telephone No. 703-308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IL01/00669

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claim Nos.: 2,3,16,17,21,22,26,27,34,35,42 and 43
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
The claims encompassed sequences not searchable because no CRF. See ERROR REPORT being sent along with this report.

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

PCT/IL01/00669

Continuation of B. FIELDS SEARCHED Item 3:
STN- MEDLINE, WEST- USPAT, PG Pub, and EPO, JPO, and Derwent
virus, replicon, reporter, infectious clone, drug, detecting, HCV, picornavirus, flavivirus